

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2003, 07:54:40 ; Search time 38 Seconds

(without alignments)
2287.245 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769
Sequence: 1 MSEGDAKVCVRVRPLQRE.....QAEWWTEAKKETAPCKTS 2954

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCrus_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/Backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1148.5	7.8	3248	1 US-08-353-700-1	Sequence 1, Appl1
2	1148.5	7.8	3248	5 PCT-US95-16216-1	Sequence 1, Appl1
3	1055.5	7.1	2482	1 US-08-328-254-6	Sequence 6, Appl1
4	1051	7.1	1388	4 US-09-572-191-2	Sequence 2, Appl1
5	1051	7.1	1388	4 US-09-723-262-2	Sequence 2, Appl1
6	1051	7.1	1388	4 US-09-723-219-2	Sequence 2, Appl1
7	953	6.5	1234	4 US-09-592-054-8	Sequence 8, Appl1
8	945.5	6.4	1232	4 US-09-592-054-2	Sequence 2, Appl1
9	906	6.1	1375	4 US-09-722-139-2	Sequence 2, Appl1
10	906	6.1	1375	4 US-09-721-832-2	Sequence 2, Appl1
11	906	6.1	1375	4 US-09-721-689-2	Sequence 2, Appl1
12	866	5.9	10182	4 US-09-134-001C-3159	Sequence 2, Appl1
13	841.5	5.7	1279	4 US-09-724-517-2	Sequence 2, Appl1
14	841.5	5.7	1279	4 US-09-641-807A-2	Sequence 2, Appl1
15	841.5	5.7	1279	4 US-09-723-096-2	Sequence 2, Appl1
16	769	5.2	3696	4 US-09-134-001C-5080	Sequence 5080, Ap
17	763	5.2	955	2 US-08-428-414A-3	Sequence 3, Appl1
18	759	5.1	955	1 US-08-006-676B-1	Sequence 1, Appl1
19	759	5.1	955	1 US-08-282-845-2	Sequence 1, Appl1
20	759	5.1	955	5 PCT-US94-00324-1	Sequence 8, Appl1
21	758.5	5.1	1066	4 US-09-541-782-8	Sequence 8, Appl1
22	758.5	5.1	1066	4 US-09-723-820-8	Sequence 2, Appl1
23	756.5	5.1	1637	4 US-09-718-692-2	Sequence 2, Appl1
24	756.5	5.1	1637	4 US-09-718-852-2	Sequence 2, Appl1
25	756.5	5.1	1637	4 US-09-718-815-2	Sequence 2, Appl1
26	748.5	5.1	2101	1 US-08-195-487-4	Sequence 4, Appl1
27	748.5	5.1	2101	5 PCT-US93-06160-4	Sequence 4, Appl1

28	744.5	5.0	2101	1 US-08-466-390-4	Sequence 4, Appl1
29	744.5	5.0	2101	1 US-08-470-950-4	Sequence 4, Appl1
30	744.5	5.0	2101	1 US-08-467-781-4	Sequence 4, Appl1
31	744.5	5.0	2101	2 US-08-483-924-4	Sequence 4, Appl1
32	744.5	5.0	2101	4 US-09-452-294-1	Sequence 1, Appl1
33	733	5.0	1073	4 US-09-541-782-6	Sequence 6, Appl1
34	733	5.0	1073	4 US-09-723-820-6	Sequence 6, Appl1
35	723	4.9	473	4 US-09-592-054-6	Sequence 6, Appl1
36	723	4.9	522	4 US-09-592-034-4	Sequence 6, Appl1
37	717.5	4.9	1898	1 US-08-056-200-94	Sequence 4, Appl1
38	717.5	4.9	1898	2 US-08-800-644-94	Sequence 94, Appl1
39	704	4.8	1184	4 US-09-541-782-2	Sequence 2, Appl1
40	704	4.8	1184	4 US-09-723-820-2	Sequence 2, Appl1
41	702	4.8	1057	4 US-09-541-782-10	Sequence 10, Appl1
42	702	4.8	1057	4 US-09-723-820-10	Sequence 10, Appl1
43	696	4.7	411	2 US-08-713-815A-4	Sequence 4, Appl1
44	696	4.7	441	2 US-08-713-815A-3	Sequence 3, Appl1
45	692	4.7	1053	4 US-09-724-519-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-353-700-1
Sequence 1, Application US/08353700

Patent No. 5599919

GENERAL INFORMATION:

APPLICANT: YEN, TIMOTHY J.

APPLICANT: RATNER, JEROME B.

TITLE OF INVENTION: NUCLEIC ACID ENCODING A

TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCORE PROTEIN,

TITLE OF INVENTION: AND METHODS OF USE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: DANN, DOREMAN, HERRELL AND SKILLMAN

STREET: 1601 MARKET STREET, SUITE 720

CITY: PHILADELPHIA

STATE: PA

COUNTRY: USA

ZIP: 19103-2307

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,700

FILING DATE: 09-DEC-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: REED, JANET E.

REGISTRATION NUMBER: 36,252

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 563-4100

TELEFAX: (215) 563-4044

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3248 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: HUMAN

US-08-353-700-1

Query Match 7.8%; Score 1148.5; DB 1; Length 3248;
Best Local Similarity 20.3%; Pred. No. 2.3e-48;
Matches 715; Conservative 626; Mismatches 1180; Indels 995; Gaps 139;

QY 106 QAIQIE---VKIIOEIPNREFLLRVSMEITVEYWKOLICDRKKPLREIDERNRY 161
 Db 18 OKIOELEGOLDKIKKEKQROFOLDLEAPOKOTOK---VENEKTEGMLKRENDRIAE 74
 QY 162 VADITEELVAVPEHVIOMIKKEKERNHYGETKMANDHSSRSHTIFRMIVESRDNNDPTNSE 221
 Db 75 ICESLETKOKISHELO-VKESQYN--FOEGOLNSGKKOLEKLEQEL--KRCSELESERQ 129
 QY 222 NCDGAVVSHLN-----LVDLASERASOTGABGVRLKEGCINNSRLFLGOV--- 269
 Db 130 QAAQSAADVS--LNPENTPQKIFTEPPLTPSOYSGSKYEDLEKYNKEVEERKRLAEAEKAL 188
 QY 270 -IKLSGOGAGFINYDSDKILRLONS-----LGNATVYICITTPPSFD 315
 Db 189 QAKKASOTLPQATNHRD--TARHOASSVPSWQOEKTPSHLSSNQR-----TPIRRD 240
 QY 316 -----ETLSILOFA-----STAKHVNTPHNEVLDDEALIKRYREILDLK 357
 Db 241 FSASYFSGELEVPTRSTLOIGKRDANSFPNGNSSPHLLDOL--KAQNELRKINIELE 298
 QY 358 KOLENLESSSE-----TKAQAMAKEHTOLLAETIKOLHKEREDRIWHNLNIYVASSQES 411
 Db 299 LRLOGHEKEMKGYNKFOEIOLOLEKAKVELIEKEKVLNCRDELVTTAQYDQASTKYT 358
 QY 412 QODQRYVR-----KRYVTWAPGKIONSLHAGVSDFDMLSRLPGNF-----SKR 455
 Db 359 ALBOKIKKLTLDLSCORONASARCISLEFORIKKEKEFEQELSRQOSFOTLOEOCIOMK 418
 QY 456 AKFS--DMSFPEIDSVCTEFSDFDALSMDSNGIDAENML-----ASKYTHREKSTLH 509
 Db 419 ARLOEOLOAKNMHNVLQAEIDKLTYSKQOJLENNLEEFKOKLRAEOAFQASQIKRELR 478
 QY 510 QSMIDFQISDSYOFHDSKRENOLOYLPKDSGMAECHKAFEREKITSILOQOSKEEEK 569
 Db 479 KSMERKKNMLKSHSQAKAREVCHLEAEKLNKICOLNOS-----QNFMEEMAKMTSQ 533
 QY 570 KELVOSFEIKTAELEQISVAKNLEMYTNSREHSINAEVOTDYKEVYKREMSYVLGDSG 629
 Db 534 ETMLRDLOEKINOQENSLTLEKILAVA-----DLEKORDOSODILKRR-----EHN 580
 QY 630 YNASDSDLODSSVQGRKLSSSHDECIERHMKLEOKIYDLEFIEIENLK-----KSENDK-- 683
 Db 581 IEQULDKISTEKEKSKALISA-----LELKKYTEBELKEKILFSCWKSENEKL 630
 QY 684 -QKSEODFMES---IOLC-----EALMAKANALEALMRDNFNIT 723
 Db 631 TQMESEKENLQSKINHLETCLKTOQIKSHEYNERVRLMEMRENLSVEIRMLHNVLDSK 690
 QY 724 LENETLKREIADLERS-----LKENOTNEFEIIEKETOKEHEKOLNHE 767
 Db 691 VEVEFOKLAUVELQOKAFESDQKHQKEIENMLKTSOLTGOVEDL-----EHKLOJLS- 743
 QY 768 IGLSKLIYENAMYNONLEEDLETKTLKKEOJOL----- 803
 Db 744 ---NEIMDKORCY--QDLHAETESLRLDKLSKDSASLVNEDHQRSLAFDQOPAMHNSFA 798
 QY 804 -----AELRRKADN-----LOKRYNFDLSVSM-----GDSEKICEEITFO 838
 Db 799 NIIEGOGSMPSESRSECRLEADQSPKNSAILQKRVDSLEFSLESQKQMSNDLOKQCEELVQ 858
 QY 839 LKQSLSDAVALYTRDAQKECSPLRSLENLEKEMEDTSMNYN-----QKEKAASLAF 888
 Db 859 IKGEIE--ENLMAKQDMQOSFYAETSORISKLOEDTSAHONVAVETLSALENKEKEJOLL 916
 QY 889 EKOLETEKSNYKKME-----ADLOKELO-----SAFNINYNLN 921
 Db 917 NDKVETEOAEIOBLKKSNNLDESLAKEJOLLSETLSLEKEMSSITILNKRIEELTQEN 976
 QY 922 GLL-----AGRYPRDLSRV-----ELEKVSSESKOLE----- 950
 Db 977 GTLKEINASINQEKMLIQKSESFANYIDEREKSSISELSDYKQOKELIILLORCSEETGNAY 1036
 QY 951 -----KALEEKNA-----LENEVYCLSYKFLPNEVECLKNOIKASSEIML-LKQSE 998

Db 1037 EDLSQKYKQAQKRNKLBELINECTSLCENR--KNELBOLKEAFKHOEFELTKIAFAEE 1094
 QY 999 HSASITISKOELTMO-----EOSE-----QILOLDETHTQOSVY---OOTE 1036
 Db 1095 RNQNLMLETELEYOQALRSEMTDNONNSKSEAGLKOETITWTLKEONKQKEVNDLLOENE 1154
 QY 1037 EGYLEKKKHH--DLFEKYIRNKSEAEDLLREMNILKGTME-----SVE 1078
 Db 1155 QLMKYVKTTHQONLESEPIRNSVYKERESERNQCNPKFQMDLEVEKIEISDYSVNAQVOLE 1214
 QY 1079 VAIADTKHELEFTIRDKEOLEHKKYFFQAMOTIPPTPLSLSPSKVLVENSODPLEI 1138
 Db 1215 AMLRNKELKLOESEKEKECLOHE-----LOTI-----RDLFTSNLODMQSOEISGL 1261
 QY 1139 NDY-----HNLALATERNNIIVWCLETERNSLKFEQVIDLNLOLSAQOSIE- 1185
 Db 1262 KCEIDAEEKYISGPHELTSONDANHLQCSLOTTWKNLNE---LEKICEILOAKVEL 1317
 QY 1186 -----KSDLOKPKODLEGEVYKLLLEMLLKG----- 1212
 Db 1318 VTELNDSRBCITATRKMAEYVGKILNVEKILINDSGILHGLVEDIPGEGEFOPNQEH 1377
 QY 1213 -----HTDSOLSTEKLOENLVEYTEKLOTOLOEMKNITTEIRNELQTNED 1258
 Db 1378 PVSIALPDESNSYEHLT---LSDEVOUMHFAELOEFSLQSHKILHDQOMSKSE 1434
 QY 1259 LKAHDSLKODLSENIQSIETQDELRAAQEELREOKOLVDSFPROLDQSVGISPNHD 1318
 Db 1435 LQTYVDSLK--AENVLVS---TNLRNFGDLVKEMQ--GLEBLVPLSLSSQVPPDS 1485
 QY 1319 AVANOEKVSIGEVNSIOSEMLRGERDELQSCALYSELE-----LNA 1362
 Db 1486 SLVS-----SLGD--SSEYRALLE-----QGDMSLTSNLEGAVSANQCVDEYFCSLOT 1533
 QY 1363 HVSVEGENEITRKNGLEKEILKSE--ESEVLKSMLENKEDNNKL----- 1410
 Db 1534 YVDSIKRKNLYLTNLNRNFGDLVKEMQGLEGLVPLSSQVPPDSSLSLGDSSFYR 1593
 QY 1411 ---EQAEYS-----SKENQPSLEEVY-----SGSQKINDEI---E 1440
 Db 1594 ALLEQTDMSLTSNLEGAVSANQCVDEYFCSLOEBENLTRKETPPAPAKGVEELSICE 1653
 QY 1441 VIKRAQKAAEERLE---IDRDYFELVQATNLYE--GKLEPPLQADHE--EDSIDRS 1493
 Db 1654 VYROSLEKLEKESOGCINKKETOELBOJLSSBROELDCLROYLSENDQOMOKLTSYV 1713
 QY 1494 EEMEIKVIGERLEKERNQYLE---RLOEKLELSNKLLEILOKEMETSVLKDDLOOKLES 1549
 Db 1714 LEMSKILAARKKQTOEOLSTLEVARLOGLDLSR-----SLGIDTPEADIOG 1762
 QY 1550 LLSBNITLKENIDITTLKHNHDQAOLOK-----TQOELQOL--AKNULATAASD 1594
 Db 1763 -----RNSCDSIKSEHTSETTERPKHDVHOICDKRAQOIDLNLTEKITERGAAPKPG 1815
 QY 1595 NCPITQOKET-----SADCVHPL-----EEKILLTEELHOK 1626
 Db 1816 ECSEGEQSPDNIYEPGDEKTOQSSECSISELSPGPNALYPMDFLGNEDLHNLOLQRYKET 1875
 QY 1627 TNEOEKILHEKNELEQAOVELKCEVHLMKSMTESKSLDESLOHEKHDTBOQLA----- 1681
 Db 1876 SNEMLRILHVIJEDDR-----KVESLILNEMKELDSKL-----HLOEVOULTKTEIAC 1921
 QY 1682 --LQAOQVYTOEKKELQOHTHEHTLAVDHLKENIEL--GLNFRKNEAO--OKTKREOCLLN 1736
 Db 1922 IELERKIVGELKKNESDISEKLEYFSCDHQOELLQOVERTSEGNDSLDEHNADKSSREDIGDN 1981
 QY 1737 ENK-----ELEQSOHRLQCE-----IEELMSLKDKESAELETLESQKY 1776
 Db 1982 VAKYNDKSKERFLDYVENELISIRSEKASISIEHAYILADLEEVQYTEKLCLEKDNENKQY 2041
 QY 1777 I-NLNOEMENYMLMEELKINSQRTVIAER---DQLODULRESVMSJET----- 1821

Dh	2042	IYVLEELSVYTSBRNQDLREBDTMSKKTATLADQSEKKKERTQ-ELSHOSECHICQV	2100
QY	1822	-QDDLRKAQEAALQOQKDRVOBELTSQISVLOEKTISLENOMLYNATVKETLSERDLNQS	1880
Dh	2101	AAEAKEKTELTQTLSSDVSELLDKTHQELQSLQESQALSLTKCELENOIQLNKE	2160
QY	1881	KOHLESELETLSLKEKEP-----ALEQAKKDADAARTIDTETISNI	1920
Dh	2161	KELLVKESESQANLSSDYETKLNSKALAEALVGEFALRSSTQOEVRHQRGIEKL	2220
QY	1927	-----EEOLQOATMYKEETIYERESLSTQCKEQOLALNTEHRETLKSDQLAGMEQER	1979
Dh	2221	RVRLEADEKQDLHAELKTKERENSNL---KQVLENRELOMSENEELVI--LDQEN	2275
QY	1980	DEAANKVIALTEKMSLSEEDINENVTTLKEG-----GEKETFYLRPSQOOSOME	2032
Dh	2276	SKA-----EVEETLKTOIEEMANSLSKTFEIDLTVLRSEKENITLTKOIEQOGOLSTSD	2366
QY	2033	ELRESITKQDQLEAEKKEISEAT-----NEIKNLTKASSL--BEETIQ--NASTLN	2081
Dh	2327	KLSSFSFLSEEKQOABITQIESKSTAVEMLOQLELMEVAALACQDEIMKATQESLD	2386
QY	2082	EAVSERENLRSHKQOULVSELE-----QLSTLSPDHAPASQREKDEAVNITASLAEI	2136
Dh	2387	PRIEEHQLRNSTIEKTLARLEADEKQOLCYQOLK-----BESHADLLKGVENYLEEL	2441
QY	2137	KILTKEND---EFRRSK---ESLQOSHSLSBELCTYKTELOMKOQKEDINNKLAERV	2189
Dh	2442	ETARNQEHAALEANSKGVEETLKAKIEGMOQSLGLEDVYVTRSEKENITNELNOEQ	2501
QY	2190	KENVELLQHLSLKEQJDOJQMLERNEKLRNYLTCMKMIDKEKLSVLRKMONEQOED	2249
Dh	2502	ERISL-----ELINSFEN-----LLOEKQOEKVOMKE	2530
QY	2250	DVAERMDILESRAOIEQIEMEKISAVYSEQHTLS---SLSELOKETEPANHCYLTNKE	2306
Dh	2531	KSSITAMEKLOT--QKLEMBERYAALHNOBACAKAEQULSSQVE-----CL-----E	2575
QY	2307	SLSTLSRSFQSLQTEHVKLNTQLTLNLFKVVYTAUVAKEDHSLINDYERKDLAAEQR	2366
Dh	2576	LEKQOLLQGDDEAKNNYIYQSSVKGHLQD-----VEDQKQLEKKEQDE	2619
QY	2367	HDELRLQLOCLEQGRKWSASABEELK-----CEIEFLNELLFKKANTIQSVODFSEVQ	2422
Dh	2620	ISRKKNQIOQEOELVSLQSVGEHOLMKQONLELNLVLEKQITQVQSSNASTLOTL	2679
QY	2423	VFLNQVSTLOEELLE--HKKGFOMOLEEFG-----DLHYARK---LSEMOQEN	2467
Dh	2680	EVLQSSKTNLENELFKMKMSFVEKVKTKAPKATELOREHNEAOKTAELOEELSEK	2739
QY	2468	RRIASTIQLTLKRLKAVQSKIOREITVYVLNQEAKLOEKKQONELKMRHEHGPSASY	2527
Dh	2740	NRLAGEHOLLLEIKS---SKQD-----LKELTLENSLELK-----SLDC	2776
QY	2528	MEENMALLGLTKTVQDESKKLOSRIKMLENELNLVKDAMKGGK--VALIQDKLSRN	2585
Dh	2777	MHKQVEKEG---KVREIELAYOLR-----LHEAKKQHQALLTD--TKMQ	2816
QY	2586	AEAEELNMOYKLRKQONLOAMKELENLQKNVAKGAVPYKREIDN-----LTKTVK	2638
Dh	2817	YEVEIQTYREKLTSKESCLSSQKLELIDLKSS-----KELLNNSLKATQIILEELK	2868
QY	2639	IEMEKIY-----SKATQOELALKSC--LEDEEGIRLRYELRLRAADNDVYCV	2688
Dh	2869	TKMNLKTVYQNLKENNRAGKKMLIKQKQLOEBEKETELQKLELSQLODAQEKOT	2924
QY	2689	PKDYQKASTPPTYCGGSGSIVQSTAMLYQSE---KALE---BELSHYKKKYHHHSRT	2741
Dh	2925	-----GYVMOTKYVDELTTEIKETLKEKTEKDEAYLDKCYSL---	2963
QY	2742	MSSEDDKKTAKKADSHSHTQSSH-----RQSPKHTETYYRGVTPYRSEMPSL---	2791
Dh	2964	LISHEKLEKKEMLETOVNAHLCQSQKODSRSP-----LLGVPVPSPPSPVSTYETKL	3017

Db 75 IESLEKTKOKISHELQ-VKESOVN--FOEGOLNSGKOIEKLEOEL--KRCSELESQ 129
 QY 222 NCDGAVVSHLN-----LVDLASERASOTGAEGVRLKEGNINSLEFILOV--- 269
 Db 130 QAAQAGADVY-LNRCPTPKOIFITPTPLTPSOYIGSGKIEDKEKYNKEVERKRLAEVAL 188
 QY 270 -IKKISDGOAGFIVYRDSKILTRILONS-----LGNKATVIYICTITPVSPD 315
 Db 189 QAKKASOTLPQATMHRD--IARHQAASSVFSMOOEKTPSHLSNSOR-----TPIRD 240
 QY 316 -----ETLSTLOFA-----STAKHVNTPIHNEVLDEBALLKRYKRIIDLK 357
 Db 241 FSAVFSGELTVPSTSLQIGKRANSFSGNSSPHILDOL--KAONOELRNKINLE 298
 QY 358 KOLENLESSE-----TKQAAMAKEHTLOLAETIKOLHEKREDRIWHLJTNIVASOES 411
 Db 239 IRLQHEKEMKGVNKFQELQOLEKAVELIEKKEVYLNKCDDELVRTTAOYDAQSTYIT 358
 QY 412 QODQVYK-----KRRVYAPKOTONSILHASGVSPDMLSRLPGNF-----SKR 455
 Db 359 ALEQIKLKLTEDLSCORONAESARCSLEOKIKEKEFEQELSRQORSFOTLDQECIQMK 418
 QY 456 AKFS-DMSFPFIDSVCTEFSDFDALSMDNSGIDAEMNL-----ASXYTHREKISLH 509
 Db 419 ARTLOEOAKRMHNVLAELDKLTSVQOLENNEEFKOKILCRAGQAFPOASQIKENELR 478
 QY 510 QSMIDEGOISDSVOPHSSKENQOLYLPKDSGDMAECKRASFKEITSLQOOLOSKEEK 569
 Db 479 RSMEMKKNMLKSHSPQKAREVCHLEALKNIKOCINOS-----QNFABEMKAKNQS 533
 QY 570 KELVOSFLKIALEEOISVYAKNLEMTNREHINAEVOTDVEKVEYKREMSVYLDGSG 629
 Db 534 ETMLDLQEKINQOONSJLTLEKLKLAVA-----DLEKRDQSODILKRR-----EHN 580
 QY 630 YNASNSDLQDSDVQKRLSSHDECIEHRKMLEOKIYDLFEITENLK-----KSENDK-- 683
 Db 581 IEQLNDKISTEKESKALLSA-----LELKKEYEELKEKELTFCQWKSENEKEL 630
 QY 684 -OKSEODFMS-----IOLC-----EAIMAKANALELALMRDNFNIT 723
 Db 631 TOMSEKKNLQSKINHTECLTKTOOIKSHEYNERYRLIEMREMI,SVLEIRMLHNVLLSKS 690
 QY 724 LENETLKREIADLERS-----LKENOFTNEFELLEKTEKOEHEAOILHE 767
 Db 691 VEVELOKAVYELQOKAFSPQKHOKETENMCILTSQITQVEYDL-----EHLQOLLS- 743
 QY 768 IGSILKLVENAEKYNLEEDLETKKILKEQETOL----- 803
 Db 744 -----NEIMDKDRCY--QDLHAEYESLRDLKSKDASLVYTNEDHQSILAFDQOPAMHHSFA 798
 QY 804 -----ALERRADN-----LOKVRNFDLSVSM-----GSEKICEBIFQ 838
 Db 799 NITEQOSMSESEKRELEADOSPKNASAILONRVDSLEFSESOKONNSDLOKOCCELYVQ 858
 QY 839 LKOSLSDAEAVYTRDAQECSEFLRSENLELKEKEMEDTSMWY-----OKEKASLFF 888
 Db 859 IKGRIE--ENIMKRAEQMHQSFVAETISQIRISKLQEDTSAHONVVAETISALENKEKEIQL 916
 QY 889 EKQLETKSNYKKE-----ADLOKELQ-----SAFNEIYLN 921
 Db 917 NDKVETBOAELOELKSNHLEEDSLKELQILSETLSLEKKEMSSIILNKRETEBELQEN 976
 QY 922 GLI-----AGVYRDLRLSHV-----ELEKKVSEFSQOLE----- 950
 Db 977 GYTLKETIASLNOEKMNLIQSESEFANTYDERKISITSELSDQYOEKILLILOKCEETGNAY 1036
 QY 951 -----KALEKRNA-----LENEVYCLSEKFLPNEVECKNOISKASEEIML--LKOEGE 998
 Db 1037 EDLSQKTKAAQEKNSKLECLNBECTSLCENR--KNELEQJKEAFKAFKHOEFLKLAFAE 1094
 QY 999 HSAIISIKOIIIMO-----EQSE-----QILOLDEVYTHQSKV-----QOTE 1036
 Db 1095 RNONLMELETVOOALRESEMTDNQONNSKSEAGIKOIEIMTLKEEONKMEQVANDLLQENE 1154

QY 1037 EOYLEMKRMD--DLFEKTYINKSEADLLKEMENLKTME-----SVE 1078
 Db 1155 QLMKMYKTRKHCQMLNESPINNNSVKERESEKRNOCNFKROMDLVEKELSLDSYNALVOLFE 1214
 QY 1079 VKIADTKHELEETITRDKBOLHEKKYFFQAMQITFPTPLSDS,LPKPKLYVGNQODTEI 1138
 Db 1215 AMLNKKELKLOESBEKREKLOHE-----LOTI-----RGDELTSLDQMSQOELISGL 1261
 QY 1139 NDY-----HNILALATERNNIMVOCLETFRNSLKEQVIDLNTQLOSLQASIE- 1185
 Db 1262 KDCEIDAEEKYISGPHLESTSQONDAHLQCSIQTMKNLE-----LEKICEIILQAEKVEL 1317
 QY 1186 -----KSDLOKPRODSEGEVKLLMEMLK----- 1212
 Db 1318 VTELNDSRSECTITRKMAEVEYKGLKNEVKTLINDSDGLNGELVEDIPGEBFGEQONEQ 1377
 QY 1213 -----HLTDQSLSIEKLOLENLEVTEKLOTLQOEEMKNITIERNELQTFED 1258
 Db 1378 FVSLAPIDESNSYBHLT-----LSDEKVOHFAELQEKFLSLQSEHKILHDQCMSSKME 1434
 QY 1259 LKAHDSLSKQDLSINIQOSITQODELPAQOELREQOYVDSFRQOLLDCSVGISSEPHD 1318
 Db 1435 LQYVYDSLK--ABNLVLS--TNLRNFOGLYKEMQOL--GLEEGIVPSLSSQVYDSS 1485
 QY 1319 AVANOEKYSIGEVSLSQSEMLRGERDELQTSKALVSELE-----LLRA 1362
 Db 1486 SLSS-----SLGD--SFYRALLE-----QTDGMSLSLNLGAVANSANQSVDEVPCCSLQOT 1533
 QY 1363 HVSVEGENLEITKNGLEKEILKSE--ESEVYKSMLENEKEDNNKLR----- 1410
 Db 1534 YVDSLKAENLVLSYNLRNFGDLYKEMQOLGEBIVYLSSSQVYDSSSLSDSGFYR 1593
 QY 1411 --BOAEYS-----SKENQFSLEEV-----GSSQULYDEI--E 1440
 Db 1594 ALBQOTDMLSLSLSEVAVANQSVDEVFCSSLOEBENLTKRTPSPAPAGVLELSLCE 1653
 QY 1441 VLAQOLKAAEBERLE--IKRDYFELVQYANTNLYE--GKLETPLOADHE--EDSIDRRS 1493
 Db 1654 VYROSLEKLEKEMSGQIMKKEKIEOELQOLLSERQDLQTRQOYISENMOQOKILSVY 1713
 QY 1494 EMEIKVYLGKLEKRNQVILE--RQOEKLELSKNLEIIOKEMESVYLLKDLQOKLES 1549
 Db 1714 LEMESKILAAKKQOTQOLSTELEVARLOLOGLDLSR-----SLGIDTEBALQ 1762
 QY 1550 LLSENILKENIDITLKHSDTOAOLQ-----TQOEIOL-----AKNULATAASD 1594
 Db 1763 -----RNSCDSIKENHSETTERTPKHDVHOICQKDAQODLNLIEKITTEGAVKPTG 1815
 QY 1595 NCPITQOKET-----SADCVHPL-----EKKILLITTEELHOK 1626
 Db 1816 ECGEQSPDITNYPPEDEKTOQSSEICISELSFGPNALVPMDFLQNOEDIHNLQVKEET 1875
 QY 1627 TNEOEKLEKKNLEBOAQVELKCEVEHLKMSIESSSLSLELOHEKHPBOOLLA----- 1681
 Db 1876 SBNENLRHLVIEBDR-----KVESLNLNEMKELOSKL-----HLOEVDLMTKIEAC 1921
 QY 1682 --LKQOQVYVTOEKEKLOOTHEHLTAEVDLKENIEL--GLNFKNBAQ--OKTTKEOCLIN 1736
 Db 1922 IELEKIVGELKKNNSDLSEKLEVFSCDHQELLQRYVTSGLNSDLEMHMADKSRREDIGN 1981
 QY 1737 ENK-----ELBOSORLOCE-----IEELMKSLKDKESALETLKSEQRY 1776
 Db 1982 VAVVNDSMKRELFYDENELSRIRSEKASIEHREALYLEADLEAVQOETKLERDNEKQRY 2041
 QY 1777 I--NLNOEMEMVLEMELEKNSORTVYABR--DOLODD,RESVENSIEI----- 1821
 Db 2042 IYCLEBELSVYSEKQOLGELDITMSKKTALDQOSSEKKEKTO--ELBHSQSECLHCIOV 2100
 QY 1822 -ODDLKRAQOALQOQKQVQELTSQISVLOEKISLLENOMLYNVAIVKETLSEBDLNS 1880
 Db 2101 AEADEVKTELLQTLSSDVSEILKDKTHQEKLOLSLEKXSQALSILTKCBLENGQINQANKE 2160


```

Db 163 HSFANIIGEGSPMSESECRLEADOSPKNSAIIONRVDSLEFLESQOKOM-----SD 216
QY 756 TOEHEAOLIHIGSLKLKLYENAMNTONIBEDLETYKLLKEOIOALERRAAD--- 812
Db 217 LOKOCE-ELVOIGKEIEENIMKAEOMHOSFV--AETSORISKLOEDTSQAONVAEFLISA 273
QY 813 LOKKVRNFOLSVSGDSEKICEEFIFOLKOSLSAENAVTRAOKCSPILRESENELEKKE 872
Db 274 LENNEKLOLINDKVEEO--AETIOELKKS---NHLEDSLEKLOTL-SETISLEKKE 326
QY 873 DTSNMVOKERKAASLEPEKOLETERKSNNKKAADLOKE---LOSAPENIYVINGLAKGY 928
Db 327 SSISLAKRE-----IEELTOENGTLKEINASLNECKMULIOKSESFANYID----- 373
QY 929 PROLLSVELEKVSSEKOLE-----KALEKKA-----LE 960
Db 374 -----ERKESISELSDYKOEKILLLORCETGNAYDLISOKYAAOEKSKLECL 425
QY 961 NEVYCLSEYKFLPREVECLKNQISKASEEIML-LKGEHSASTISKOETIMQ----- 1012
Db 426 NECTSLCENR--KNELQOLKAPAKHEQEFITKLAFABERNONIMLEIYQOALRESEMT 483
QY 1013 -----EOSE-----QILOLDEVYHTOSKV---QOTEBOYLEKRMHD--DLEFYIR 1055
Db 484 DNQNSKSEAGGLKOEIMTLKEQNKMKQKVEYNDILOENEOIMKMKYKHECQONLESEPIR 543
QY 1056 NKSABELLREMENTKCTME-----SVEKVIADTKHELETTRDKCOLL 1099
Db 544 NSYKERSERNOCKFQOMOLEYKELSDYSTNAOLVOLLEAMLRKELKLOESEKELQ 603
QY 1100 HEKKYFQOAMOTFPIPLSDSLPSKLVGNSQDPIEINDY-----HNILAL 1147
Db 604 HE-----LOTI-----RGDETSNLODMQSOEISGLKDCELDAREKYSIGHELSTS 650
QY 1148 ATERNNIMVCLFEBRNSLKEQVIDLNTOLOSLAKOSTE-----KSDLOKQODLEEG 1199
Db 651 QNDNANHLOCSQOTTKMLNE---LEKICETLOAKELYELTDLNDSRSECTTATRKAAEE 706
QY 1200 EVKLLLEMELKLG-----LEKICETLOAKELYELTDLNDSRSECTTATRKAAEE 706
Db 707 VGKLLNEVKILINDSGILHGEIVEDIGGEGEOPNEOHVSLAPLDESNSTYELT---L 763
QY 1220 SIEKLOLENEVTEKLOILOEMKNITIERNELOTFNEEDIKAEHDSLKODLSENIEOSIE 1279
Db 764 SDRKVOHMFALOEKFLSLQSEHKILHDHCQMSKMSLEQTYVDSLK---AENLVLS-- 818
QY 1280 TODBLRAOEBLROKOLVDSFRQOLLDCSVGISSPNHDAVANOEKSLGSEVNSLOSSEM 1339
Db 819 --TNLRNFQODLVKEMQL--GLEEGVPSLSSSCVPPSSSLS---SLGD-SSFYRAL 868
QY 1340 RGEDELOQSCALVSELE--LLRAHVKSEGENLEITRKILNGLEKEILIGKSEES----- 1392
Db 869 E-----QIGMSLISLNEGVANSQOSVD---EVF--CSSLOEBMLTRKETPSARAK 916
QY 1393 -EYVLSKMLENIKEDNNKILKEOAEYSKENOFSLEEVSSQKLVIDEIYVLAOKLAAEE 1451
Db 917 VEELESICEFYAROSLEKLEEMESQIMKNK-----BIOLELOLSSERQ 961
QY 1452 RLEIKNDYELVOTANTNLVEGKLETPLOADHEDSIDRSEMEKIVLGEKLERNOYL 1511
Db 962 ELDCLRVOYLS-----ENEOWOOKLTVLTEMESKLAEEKOTOLS 1003
QY 1512 LE-----RLOEKLETSNKLLEILOKEMETSVLLKDDLOOKLESLSSENIILKENIDTILKH 1567
Db 1004 LELEVALLOLOGLDLSR-----SLGIDTEDIAIOG-----RRESCDISKEH 1045
QY 1568 HSDTOALOKR-----TOELOL-----AKNLAIASDNCPIQOETK----- 1604
Db 1046 TSEETERTPRHDVNOQIDKAOQODLNDIEKITETGALKPGECSGSGSPDTNPERPED 1105
QY 1605 ----SADCVHPL-----EKKILLTEELHOKTNEOKLLHEKNELBOAO 1644
Db 1106 KTQSSSEICISLSPSGNALVPMDFLGNQEDIHMLQLRVRETSNENRLLTHVIEDRDR-- 1163

```

```

QY 1645 VELKCEVHEKMMISKSSLESLOHEKHDTEOULLA-----LKQOQVYTOEKKELQ 1697
Db 1164 ----EYESLLENKEKELDSKL-----HIOEQOLMTKIEACLETETKIVGELKKESSDS 1211
QY 1698 QTEHHLTAEDVHLKENIEL--GLNFRNEAO-QOTTEOCOLNENK-----BLEOS 1744
Db 1212 EKLEYSCDHQELLQVETSEGLNSDLEMHADSSNEDIDONVAKVNDMSKERFLVEYN 1271
QY 1745 OHRLOCE-----IEELMSLKOKESALETLKESEOKYV--NLNOEMENYLMMEEL 1793
Db 1272 LSRIRSEKASIEHEVALYLEADLEVYOTEKLEKENDENKOKVIVCELEELSVYTSERNOL 1331
QY 1794 KNSQRPVIAER---DOLODIDRESVEKSTIPT-----ODDLRKOALQOQOKDK 1838
Db 1332 RGELODMSKRTYALDOLSEKMKERTQ-ELRSHOSECHIQVAEAYEKETELLQTLSSD 1390
QY 1839 VOELTSQISVLOERKISLENOMLYNAVYKETSERDILNOSKOHDFSEITELSLKEK 1898
Db 1391 VSELLMKOTHLQERKLOSLKEDSOALSUTKCELENQOLNKEKRELLVKESELSQARLES 1450
QY 1899 EF-----ALBOAKDKAARKTIDITEKISNI-----BEOLLOQATNL 1937
Db 1451 DYERLANSKALAEALVEKGFALRLSSTOEYVQLRGJELKLRVLEADOKLOLHAERKL 1510
QY 1938 KETLYERESLQCKEOLATHEHLRETILKSDALGKMEORDEANKVYALTEKMSLE 1997
Db 1511 KERERENDSL--KDAYENLERELOMSEENQELVI--LOANENKA-----EVETLK 1556
QY 1998 EQJNEVNTYKKEG-----GEKETFYLOPSPKQOSSOMELRBSLTKYDLOLEBAEK 2050
Db 1557 TOJEFMARSLKVELDLVTLRSEKENLTKOIOERKOGOLSLDLNLSFKFLTEKKEQAL 1616
QY 2051 EISEAT-----NETKNLTAKISSL--BEELIQ--NASLINAVSREMLRHSKOOLVS 2099
Db 1617 QIKESKTAVEMLONOUKELNEBAVALCGOETMKATGESLDPPIEBEHOLRNISTEKLA 1676
QY 2100 ELE-----QSLTSLKSRHFAOSKREKDEAVNKIASIAEIKILTKEMD---EPRDSK 2150
Db 1677 RLEADEKKOICVYLOQLK-----ESEHNDLLKGVENLEBELEIARNOHALAEVNSK 1731
QY 2151 ---ESLOOSSHLSEELCTYKTELQOMLKQOKEDINNKLAKEVVEDELLOHLSLKEQLD 2207
Db 1732 GEVETLAKLEGVOTOSLGRGELDVVYTRSEKENITJELVEOEORISL----- 1779
QY 2208 QIQMELNEKLRVYELCEKADIMEKEISVLRIMONPEQOEDVYAEEMDILLESNOEIOE 2267
Db 1780 ----EITNSSFEN-----ILOKEQOEVOKKESSTRAMELOQT---QIKE 1817
QY 2268 IMEKISAVYSEOHITLSSLSSELOKETEAKHGMNLKESLSSLSRSPGSIQOTEHVKNL 2327
Db 1818 LNERVALHNDQ-----EACK-----AKEONLSSOYVCELELEKAOULL 1854
QY 2328 TOLOTLNKKFYVYKTAAYVEDSLTKDYKDLAABOKRDE---RLQOLCILEONGRK 2384
Db 1855 QGLDEAKNNYIVLOSSV---NGLIQVEDGOKLEKKOEELSRKLNQIDODEQVLSKL 1909
QY 2385 SDSASEELK-----CEIEFLNELFKKANTIGSVODEFSVOYFVLNOVGSTLOBELE-HK 2439
Db 1910 SOYBGEHQLMKEONLEIRNLTVLEQKIQVLOSKNASLODTLEVLOSSYKNLENLEELTK 1969
QY 2440 KGFQOWLEERG-----DLHADYAK---LSEGMQOENRRIASTIQLTKRLKAVY 2485
Db 1970 MDKMSFVEYKNNKTAKETELQREHNEHAQTAELQELTESSEKNRILAGELQDLLEIKS-- 2027
QY 2486 QSKIOREITVYLAOFEAKLOEKKOEKONKELRRMHNHSPASVYBEBNARLLGLIKTYODE 2545
Db 2028 -SKDQ-----LKEITLSENELK-----SLDCNKKQOKEG---KYREE 2063
QY 2546 SKKQOSRIKMLENLYKNDAMHNGEK--VALIOLKLSRNAEALANAMOVLTTRKQDN 2603
Db 2064 IAEYOLR-----LHEAKKHQALLLD--TKQOYEVELOTYRERTLSKEEC 2106

```

```

OY 2604 LOAMKEENLOKNAVAGAVPYEEEDN-----LKTWVYLEMEKIKY-----SK 2648
Db 2107 LSSORLEIDLUKSS-----KEELNNSIKATQOILEELKTKTKMDKNLYVQKRENER 2158
OY 2649 ATDOEIAVLKSC--LEDKEGBRLKEELRRAADNDTVYCVPKKYOKASTEPYVCGGS 2706
Db 2159 AOGKMKLLKSCQOLEEKEEILLOKELSOLOAOEOKT----- 2166
OY 2707 GIVOSTAMVLQSE---KALE--RELSHYKKKYHLKRTWSSSEDEKRTKAKSDAHS 2759
Db 2197 GTVADTKVDELTEIETKELEETEKEADEYLDKCYSL--LISHKEKLEKAMELTOY 2253
OY 2760 SHTSSH-----KSPKIKETTYRHGCVYTPRSMPSL--HLGSRKSSSKRYVS--- 2808
Db 2254 AHLCSQSKODSRSP-----LLGPVPGPSPITPSYTEKRLSSQNKAKSGKRKSSIW 2307
OY 2809 -----PNRSEIYS-----OLVMS-----PGKGM 2837
Db 2308 ENGSGPPPAIPESPESKSKAKVAVSGIHPADETEGTEFPEEGLEPYVYKKGFADIPGKISP 2367
OY 2828 H-----KHILSPSKVGLH-----KKRALSPNRSEMPOTHOVISPCKTG 2864
Db 2368 YILRRTMATPSRLAOKIALSLPSLIGKENLAESSKPTAGSGRSQOKVAVQNSPVDVG 2427
OY 2865 ---LKKNLTESTLPDNLSPCKQOKQOVENLSPKGLF-----DVSKSMPTCPSPQ 2912
Db 2428 TILPEPTTKSPVNNLPEPSPDPSPEGGKRVYKGRVLVPSPKAGLSKSEMNCKVQ 2482

```

```

RESULT 4
US-09-572-191-2
; Sequence 2, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
; US-09-572-191-2

```

Query Match	7.1%;	Score 1051;	DB 4;	Length 1388;
Best Local Similarity	24.5%;	Pred. No. 5.2e-44;		
Matches 417;	Conservative 304;	Mismatches 548;	Indels 434;	Gaps 60

[illegible]

0Y	261	FINRDSKULTILIONSJAGNAKYITCTIPRV--FDLSTLQRPSTAKHYRTPHNE	338
Db	312	HVCYRDSLTFLRDLSDJAGNAKTAIIANVHGRSGEGLSTLNFQARKLLKNAAYNE	371
0Y	339	VLDRLALKRKRETLIDKLOLENTLESSSEFTKAQAAKKEHPOLLAETIKOLKREDRIM	398
Db	372	--DPOGNVSQDAEYKRLKEDLALASGQRPRESFLTRDCKKTNYMEYQ-----EAMLF	424
0Y	399	HLTNIVASSQESODORVYKRRRYTVAPEKIONSJHAGSVDPMLSLPGNFSKAKF	458
Db	425	F-----KXSEQEK-----SLIEKVTOLEDTLTKKEKF	452
0Y	459	SDMSPFPIIDSVCTEERSDPDALISMDSNGIDAENNLASKYTHREKTSJHOSMIDFOI	518
Db	453	-----IQSNKMYVYK--EDQIIRLEK--LH-----	474
0Y	519	SDSVQFHDSSKENOLQYLPKPDGDMACEKRSASEKEITSLDQOLQOSEKKEKELVQSEL	578
Db	475	-----KESNGFLPEEDRLL-----SELNRNETIQLREQIEHNPVAKYAMENHSL	520
0Y	579	KIALEBOLSTYKANKLEMVTNRSRHSJNAEVOYDVEKEYVYKRMAYLGGSGYNASNDQ	638
Db	521	R-----EENRRRL--LEPVAKOE-----	538
0Y	639	DSVYDGRSLSSHEDCEJHHRMLQOKIVDLEEFTEMLNKSENDOKSSEODFMESIQLC	698
Db	539	-----MDAQITAKLEKAFSEISGEMSKDK---NOQGFSPRAKOE	574
0Y	699	EALMAEKANALEETALMRDNFIILENETLKEKREIADLESJKEN--QETNEFEILEKETQ	757
Db	575	PCLFA-----NTEKILQALLQIOTELNNSKQEEYEEF-----KELT	609
0Y	758	KEHEKOLIHETGSLKTLVENAKMYNONLEEDLTETKRLKLEDOIAQLEK--RADNLO--	814
Db	610	KRQOLESELSLOQA-----NLWLEMLLA--TRACKROEV--SOLNKJHAETLKII	659
0Y	815	--KKVRNF-----DLSVSGDSEKJCEBIEFOLKOSLSDAEAVYTRDAQECSPLRENTL	865
Db	660	TTPTKAYQJLHSPVPKLSPEGSGFSL-----YTONSIIIDNDILNEPVPE-----	706
0Y	866	ELKEKEMETSNMYNOKKAKASLPEKQOLETPEKSNYKMHEDLOKE-----IOSAPNEINY	919
Db	707	-----MNDQADPAISEELRTYQEOBMSALQAKLDEEBHKNLKLOOHVDKLEH	752
0Y	920	INGLAGKVPRLDLSRYELEKKVSEFSKOLEKALEKNALENEV--TCLSEYKFLPNEVEG	978
Db	753	HSTQM-----QELFSSRI-----DMYKQOEBLLSOLANVLEKQLOTOYKKNFLASVND	802
0Y	979	LKNQISKASEITMLKOE-----GEHSAITSKOEITMOESEOIOLITDEVTHTOSK	1033
Db	803	LRYVLHSHADKELSEVSKLEYSFKNQEKEPNKLSEKHMVYOLODNLRLLENKLELESKAC	862
0Y	1032	VOOTEBOYLEKKMHNDLPEKUTYR--KSEAEDLREMEYNKGMSEVEVK-----IA	1082
Db	863	LQDSYDNLQELMKREIDQLSRNLOKFKKEKMETLAKSDNLMLLELEKAKRNNKLSLOE	922
0Y	1083	DTKHELEETIRDKEQLLHEK--KYFQAMQOTIPRITPLSDSLPSPKLYEGNSODPIEINDY	1141
Db	923	DKENSKEETIKLEVLAVQROKQETAKCEOMAKVQKLEESL-----	963
0Y	1142	HNLAETERNNTINVCLETERNSLKQEVYDNLNQLOSLOASITKESD--LQKPKDLEGE	1200
Db	964	-----LATER--VYSSLEKSRSDSKVADLMMQIOELRSSVCEETETDTLQKOLKIN	1016
0Y	1201	VK-----LLEMLLEKGLTSDQISLEKJOLENLEVELEKTOLOEBKMNITIERNELOTFNE	1257
Db	1017	KYNSALVDR-----ESRYVLIKQEOVIDDKETLH-----LRISSE	1054
0Y	1258	DLKXHDHSLKODLSENTQO-----SIEFOELRAQOEELREKQOLYDSFQOOLLDSY	1310
Db	1055	DI--EROMLCEIDLNAHVEDOANMLTEASKRISGSLGQASOAEIITKKAJLOEOLNK-----	1107

QY 1311 GISSPNHDVANOEKVSLGEVNSLOSSEMLRGERDELQTSKALVSELELLRAHVKSVEGE 1370
 Db 1108 -----NOKK-----EVEEQKNNEYNKKNOL-----EHMDSAAE 1137
 QY 1371 N-----LEITKKNLKELEIKLSESEVLSMLENKEDNNKKE--QAEYS 1417
 Db 1138 DPOSKPPPHQTHAKLLETQEOEI-----EDGRASKTSLEHLVTKLNEBEVKNNAELR 1193
 QY 1418 SKENQFSLEEVFSGSOKLVEIEVLAQOLKAERLEIKDYPELVOTAN-----1468
 Db 1194 MKEQREMENLRLESQOLIEKNMILQOLODIKQKENSQDNHNDQOLNKEQESIKER 1253
 QY 1469 ---TNVYEGKLETPLOADHEDSDIDRSEMEITVLEKLEKNQYLLERLOEKLISLNK 1525
 Db 1254 LAKSKIYEEMLR--MKADLEEVOSALYNNKEWECILRMTEDEVERTQTELSKAFQEREQLRSK 1311
 QY 1526 LEILOKEMETSVLLKDLQOKLESLSSEN--IILKENIDITIKHSHSTQOLOKTOELO 1583
 Db 1312 LEEMEERERTSOMEMLRKQVECLAEENKGLVGHQNL-----HOKIQYVVRLLKENVR 1365
 QY 1584 LAKMLATAASDNCPTIOEKETSA 1606
 Db 1366 LAETERTLRAEVNFLEKKEKRSSES 1388

RESULT 5
 US-09-723-262-2
 ; Sequence 2, Application US/09723262
 ; Patent No. 6379912
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
 ; FILE REFERENCE: 1017
 ; CURRENT APPLICATION NUMBER: US/09/723,262
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US 09/572,191
 ; PRIOR FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1388
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-723-262-2

Query Match 7.1%; Score 1051; DB 4; Length 1388;
 Best Local Similarity 24.5%; Pred. No. 5,2e-44;
 Matches 417; Conservative 304; Mismatches 548; Indels 434; Gaps 60;

QY 2 SEGDVAVGVVRVPLIORE---QGDQANIQKAGANTTQOVGT---KSFNDRVFNSSHE 55
 Db 22 NEBDATKIVVRIRKPRERSGADGEQ-NLCLSVLSTLRHSNDEPKFTFDHADVDVT 80
 QY 56 STSOIYQEIAPVPIRSALOGYNGTTFAYGOTSSGKTYTMAG-----TPNSLGIIPAI 108
 Db 81 TQESVYATVAKSIVSCEMNGYGTTFAYGOTSGKTFYMGSSEDNSHNRGVIPNSF 140
 QY 109 QEVYKIQETI-----PNRELLRVSYMEIYNETVADLDCDDRRKRPLEIREDFNNVYA 163
 Db 141 EYLFSLIDREKKEKAGAGKFLCKSFIEIYNEQIYDLT--DSASGLVLRHIKKGVYV 198
 QY 164 DLTEELVMPBEHVIQMIKKEKRNHYGETKMDHSSRSSTITIRMYVESDR-NDPTNSEN 222
 Db 199 GAVEOVVTSAAAYOVLVSGGMRNRVASTSMNRESSRSHAVFTTIESEKSNETYN--- 255
 QY 223 CDGAVVMSHLNVLDSGERSAQTGAEGVRLKEGCNINRSFLITIGQVTKKISDQAG--G 280
 Db 256 ---IRTSILNVLVDLAGSBRQKDTAEGMRLLKAGNINRSLSCLQGVITALTAVDNGKOR 311
 QY 281 FLYNDSKILTRLQNSLGNKATVITICTTPVS--FDELTSLQAPASTAKHVHNTPHVNE 338

Db 312 HVCTYRDSKTLFLRLDSISGNAKTALTITANVHPSRCGETSLTNFQRAKLLINKAVNE 371
 QY 339 VLDEDAALKRYREIIDLKQOLENLESSSETKQAAMKEBHQTOLLAIEIQLHKEREDRYM 398
 Db 372 --TQGVNSQLOAEVRYKRLQELAEASGCPPEPSFLTRDKKKNVMEYIQ-----EAMLF 424
 QY 399 HLNIYVASSQEQOQORVYKRRRYTWAPQKTONSLSHAGVSDFDMLSLPGNFSKAKAF 458
 Db 425 F-----KKEOEKK-----SLIEVYTOLEDTLLKKEKF 452
 QY 459 SDMPSEPIDDSVCTEFSDFDALSMDSNGIDAENNLASKYTHRREKTSHQSMIDFGI 518
 Db 453 -----IQSNKMVYKR-EDQILREK--LH-----474
 QY 519 SDVQVQHDSSKENQLOYLRLKROSGDMAECKRASEKETSLSQOLOQKEEKKELVQFEL 578
 Db 475 -----KESRGQFLHEEDQRLT---SELNRNEIOTLRQDIEHHPVAKYAMENHSL 520
 QY 579 KIALEEOLSVKAKNLEMTYNSREHSHINAEOVDVEKVEYVRKEMSVLGSQYNASNDIQ 638
 Db 521 R-----EENRRLR--LEPYKRAQE-----538
 QY 639 DSSVDGKRLSSSHDECIEHRKMLEQKIYDLEEFTEENLNKSSENDROKSSQEDFMESIOLC 698
 Db 539 -----MDAQTIARLEKAFSEISGMEKSDK---NOGFSPKAQKE 574
 QY 699 EAIMAEKANALEELALMRDNFNILLENETLRKREIADLERSUKEN-QETNEFEILEKETQ 757
 Db 575 PCLIFA-----NTEKLAQLOQIOTELNNSKOEVEE---KELT 609
 QY 758 KEHEAOHLIEIGLSKLKYVENAEMYNQNEEDLETYTKLKEOEIOLAEIR-RADNLO-- 814
 Db 610 RKRQLEFESELSLOKA-----NINLENLLA--TKACKREV--SOLNKHATLTKII 659
 QY 815 ---KAYRN---DLSVMSDSEKLCEITFQLKOSISDAEAYTRDAQKECSFLRSEN 865
 Db 660 TTPYAYOLSHRPVRLSPLEMSGFSGL---YTQNSSITDNDILNEPVPE-----706
 QY 866 ELKEKEDTSMNYNOKERAKSLFEKOLETEKSNYKMEADLOKE-----IOSAFNEINY 919
 Db 707 -----NMQAFEAISELRTVQPOMSALQAKLDEEHKMLKIQOHVYKLEH 752
 QY 920 INGLIAGVPRDLRSVLEIKVSEFSKOLEKALEKNALENEV-TCLSEYKFLPNEVEC 978
 Db 753 HSTOM-----OELFESSERI-----DMTKQOEELLSQNLVLEKOLEQETQKNDPLKSEVHD 802
 QY 979 LKNOISKASEETMLKOE-----GEHSATISKQELIIMQOSEQIILQLTDEVTHTQSD 1031
 Db 803 LRVYVSHADKELSVKLEYSEFKTNOEKEFNKLSRHHVQLOLDNLTLENEKLLIESKAC 862
 QY 1032 VQOTEBOYLEKMKHMDLFEKYIRN-KSEADLLREMENTLGTMESYEVK-----IA 1082
 Db 863 LODSDYNQOELIKFPIIDLSRLQVFKKENETLMSDNLNMLLEBAEKERNKSLQPEE 922
 QY 1083 DTKHLEETINDKBEULHEK-KYFQAMQOTIFPIPLTSDSLPSPKLYVGNSDPIEINDY 1141
 Db 923 DKENSKEIILVLAVRQEKQETAKCEQOMAKVQKLEESL-----963
 QY 1142 HNLIATATERNNIMVCTETENSLSKEOYITDNTQLOSIOASIKSD-IQPKODLEGE 1200
 Db 964 -----LATEK--VASSLEKSDSKVYADLAMOIOELRSSVCETIETITLKLKELNDIN 1016
 QY 1201 VK-----LLEMLLKGHLTDSQLSIEKLQLENTLEVTEKLOTLQEBMKNITTERNLEQTNFE 1257
 Db 1017 CKYNSALVDR-----ESRVLIKQOEVDITLDKETLR-----LRLTSE 1054
 QY 1258 DLKAEHDSIKODLSENIEQ-----SIEQDELRQAQOEELREKQOLVDFRQOOLDCSV 1310
 Db 1055 DI--ERDMCEDLHAHTEQLNMTLTKASKRHSGLLSQASQOEELTKREALIOLQHLT-----1107
 QY 1311 GISSPNHDVANOEKVSLGEVNSLOSSEMLRGERDELQTSKALVSELELLRAHVKSVEGE 1370

Db 1108 -----NOKK-----EEVEQKKNEYNFKMKQL-----EHYMSAAE 1137
QY 1371 N-----LEITRKLNGLEKEILGKSESEVILKSMLENKEDNNKLE--QAEYS 1417
Db 1138 DQSPKPRPHQHTLAKLEEQOEI-----EDGRASKTEHLVTKINEDREVNAELTR 1193
QY 1418 SKENQFLEEVFSSOKLVNDEIVLKAQKAERLEIKRDYFELYOTAN-----1468
Db 1194 MKQOLREMEMLRLESQOLIKEMMLLOGLDIDIKKOKENSQONHPDNOOLKNEQESIKER 1253
QY 1469 ---TNLEGLKLEPLQADHEDSIDRSEMEIKVLGEKLENOYLLELQOEKLELSNK 1525
Db 1254 LAKSKIYEMLK--MKADLEEVQSALENKMECLRMDEYERQTLSEKAFQEKELRSK 1311
QY 1526 LEIQKEMETSVLKDDLOOKLESLSN--IILKENIDTTLKHSPTQAOLOKTOQELQ 1583
Db 1312 LEEHYERERTSOEMEMLRQOVCELAENGLYGHQNL-----HQRIOYVYRLKKNVR 1365
QY 1584 LAKNLAIASDNCPIQEKETSA 1606
Db 1366 LAETEKLRAEVNFLEKKRSSES 1388
RESULT 6
US-09-723-219-2
; Sequence 2, Application US/09723219
; Patent No. 6391613
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6391613el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723, 219
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRF
; ORGANISM: Human
US-09-723-219-2
Query Match 7.1%; Score 1051; DB 4; Length 1388;
Best Local Similarity 24.5%; Pred. No. 5, 2e-44;
Matches 417; Conservative 304; Mismatches 548; Indels 434; Gaps 60;
QY 2 SEBDAYKVCYRVRPILORE---OGDDANLQWKAGNNTISQVDT---KSFNEPRVFSHE 55
Db 22 NEDDAIKVFVRIRPRRSGSADGEQ-NICLSVLSSTSLRHSNPEKPTTFDVAADVT 80
QY 56 STSQIOEIVLPIRSALQNGTIFAYGQTSQKTYTMAG-----TPNSLGIIPQAI 108
Db 81 TQSSVATVAKSTVESGMSQNGTIFAYGQTSQKTYTMAGPSQDFSNHKLKVIPIRSF 140
QY 109 QEVFKITQEI-----PNREFLRVSYMEIYNEVYKLLCDRRKKPILEDENRNYVA 163
Db 141 EYFLSDIREKEREKAGKSLCKSCFEIYNEQIYDILL--DSASAGLYREHIKKGFVV 198
QY 164 DLTEELMNVPEHYIQWIKKEKNRHYGETKMNDSHSHITFRIVTSRRR--NDPTSEN 222
Db 199 GAVEQVYTSAAEYIYVLSGWRNRRVASTSMNESSSHAVFTTTESKSENIYV--- 255
QY 223 CDGAVWVSHNLVLDIAGSEFASOTGAGVRLKGCNINRSFLITIGYIKKLSQOAG--G 280
Db 256 ----IRSLNLTVDIAGSEKQDTHAGCMRLKENGINRSLSCIGQYITLAVYGNKOR 311
QY 281 FINRDSKLTIRIIONSLGNAKTYIICTIPVS--FDETLSTLOFASTAKHVRTPHVE 338
Db 312 HVCYRDSKLTFLRLDLSIGNAKTAIILANVHPGSRCEGTSTLNFQARAKLIKKAAYNE 371

QY 339 VLDEBALIKRYRKRIEILDKKOLENLESSETKKAQAMAKEERTOLLAEIKOLHKEREDRIW 398
Db 372 --DTQGVNSQLOAEVYKRIKOLELAEIASQOTPRESEFLTRDKKTYMEYFO-----EAMLF 424
QY 399 HLNTIYVASSQESQODQVKKRRVYAPAKGIQNSLHAGVSDPMLSRLLGNSKRAKE 458
Db 425 F-----KKSQOEKK-----SLIEKTYQLEDLTLKKEKE 452
QY 459 SDMPSEFIEDSVCTEFSDFDALSMDNSGIDAEWNLASKVYTRREKTSLSHQMIDRGOI 518
Db 453 -----IQSNKMIYFR-EDQITLLEK--LH-----474
QY 519 SDSVQFHDSSKENOLOYLPKDSGMAECRAKSEKEITSLDOOQLOSKREEKEKELVSEFL 578
Db 475 -----KESRGGLPEEODRIL--SELNREIOTLREQIEHHPRAKYVMEHNSL 520
QY 579 KIAELEPOLSVKAKNLEKVTNSREHSTINAEVQTVDEKEVYKREMSVLGDSGYNANSOLQ 638
Db 521 R-----EENRRLRL--LEPVKAQE-----538
QY 639 DSSVDGKRLLSSHDECIEHRRMLEQKIVDLEEFLENKKSSENDKOKSSQEDFMESIQLC 698
Db 539 -----MDAQTLAKLEKAFSEISGMEKSDK---NOGCFSPKAQE 574
QY 699 EAIMAKANALEELALMRDNDNIILENETIKREIADLERLAKEN--QETNEFELLEKEYO 757
Db 575 PCLEFA-----NTEKLAQOLLQIOTELNNSKQYEEF---KELT 609
QY 758 KEHEAQLIHEIGSLKTLIVENAEYNNONLEDELFRTKLRKQEQLOLELR--RADNLO-- 814
Db 610 RKROLELESELSQLOKA-----NLNENLELA--TRACKQEV--SOLNKIHAETLKI 659
QY 815 ---KVRNF-----DISVKGDESEKICEETFOLOKOSLSDAEAVTRDAQKESFLRENTL 865
Db 660 TTPTRAYOLHSRPVPKLSPWEGSGSL---YTQNSSTLDDINDINEPVPE-----706
QY 866 ELKKEMDTSWVYQKQKASLFEKOLETEKSNKKKEADLOKE-----IQSAFNEINY 919
Db 707 -----MNOQAFPAISEELTYQEOMSALQAKIDEEHNLKLDQOHVADLKH 752
QY 920 LNLGLACKVRDLSRYELEKRYSEFSKOLEKALEENKALENEY--TCLSEKFLPNEVEEC 978
Db 753 HSTQM-----QELFSSERI-----DWTKQOEELLSQLNLEKQLOEQOTKNDFLKSEVHD 802
QY 979 LKNOISKASEEIMLKOE-----GEHSASISKQETIMQOSEQOLQITDEVYTHQSK 1031
Db 803 LRVVLHSADEKLSVYKLEYSFKNQKEKFKLSERHMHVOLQIDNRLLENKELLESKAC 862
QY 1032 VQOTEEQYLEKKMHNDLFEKYIRN--KSEAEDLLREMNKLGTMESVEVK-----IA 1082
Db 863 LODSYDNLQETMKETIDQLSNRLNLFKKNETLSLSDNNLLELAEERNNKSLQPEE 922
QY 1083 DTKHELEETTRDKQOLLEK--KYFFQAMQITFTPTLSDSLSPSKLYEGNSQDPEIENDY 1141
Db 923 DKENSSKEILKVLAEVQKQEKERAKCEQOMAKYQKLEESL-----963
QY 1142 HNLIALTERNNIMVCLTERBNSLKEBOYIDLNTQLOSLQASQISTSD--LOKPKODLEEGE 1200
Db 964 ----LATER--VYSLEKSSDSQKRYVADIMNOIOELRSSVCEKTEFTIDTLKQELDIN 1016
QY 1201 VK---LLELEMLKGLHLDLSQISTEKLQLENLVEYTEKLOTLOEBMKNTITERNLQNF 1257
Db 1017 CKTNSALVDR-----ESKVALTKQEVOLDLKEFLR-----LRLISE 1054
QY 1258 DLKAHSDLKODLSENIQO-----SIETODELRAQOEELREKQOLVDSFRQOLDCSV 1310
Db 1055 DI---ERMMLCEIDLAHATEQOLNMLTEASKKHSGLQASQOEBELTKKEALIOELOHRL-----1107
QY 1311 GISSPNHDAVANDKYSVLGEVNSLQSEMLKGEDELQTSKALVSELELLRAHYKSVGE 1370
Db 1108 -----NOKK-----EEVEQKKNEYNFKMKQL-----EHYMSAAE 1137

QY 1371 N-----LEITKLNGLEREKLEKSESEVLKSMLEKEDNNKLE--QAEYS 1417
 Db 1138 DQSPKTPPHQTHLAKLETQOEI---EDGRASKTSLEHVTKLNEDREVNAILR 1193
 QY 1418 SKENQSLSEVFSQSKYVDEIVELKQAKRAERLEIKDRDYELVOTAN-----1468
 Db 1194 MKEOLREKEMRLRSLQOLTEKNNLLOGLQDIDIKRQKENSQDNHNDQOLKNEDESIKER 1253
 QY 1469 ---TNLVEGKLETPLOADHEDSIDRSSEMEIKVLGKLEKRNQYLLERLOEKELETSK 1525
 Db 1254 LAKSKIYEMLK--MKADLEEVQSAALNKNKECHLRMTDEVRQYTLSSKAFQOEQJRSK 1311
 QY 1526 LEIIOKEMESVILKDDLOOKLESLSSEN--TIKENIDTTLKHSQTOAQLOKTOQELQ 1583
 Db 1312 LEEMEYERERTSQEMEMLRKQVECLAENGKLVGHQNL-----HOKIQYVRLKKNVR 1365
 QY 1584 LAKMLTAASDNCPIPOEKESA 1606
 Db 1366 LAETEKLRANVFLKKEKRSSES 1388

RESULT 7
 US-09-592-054-8
 ; Sequence 8, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684el. motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592,054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 1234
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-592-054-8

Query Match 6.5%; Score 953; DB 4; Length 1234;
 Best Local Similarity 27.2%; Pred. No. 2.9e-39;
 Matches 333; Conservative 220; Mismatches 406; Indels 239; Gaps 42;

QY 3 EGDAAVCVAVRPLQIRE--QGDQANIMQKAGNNTISQVGT--KSPFNPVFNESHSTSO 59
 Db 6 KGIPIVVALRCRPLVPEKISEGQMCISFVPGETQY--VVGIDKSTFYDFVDPCTGEDE 63
 QY 60 IYQELAVPIIRSLAQYNGTIRAYGQSSGKTYTMG-----TPNSLGITIPQAOIEVF 112
 Db 64 VENKAAVAPLIKGIFFKGYNATVLAAGQGTSGKTYSMGAVTADQENPVTGIIIPRVIOLEF 123
 QY 113 KIIQELPNEEFILRVSMEYMETVYKDLICDDRKKPLEIRDFENNVYVALUTELVAV 172
 Db 124 KEIDQKSDPEFLIKVSYLTIYIEEILDLCPSEKQAQINIRDPKGIKIVGLTEKTYIV 183
 QY 173 PEHVQIWMKGEKRNHRYGPTKANDHSSRSHITFRMIVESRDNDPTNSENCQAVVSHL 232
 Db 184 ALDVTASCLQGNNSRFTVASTANSSQSRSHALFTISLDEGKSKDNSS-----FRSL 236
 QY 233 NIVDLAGSRASQGTAEGRLEKGCNINSLILGOVTKLSLSDGAGGFINTRDSKLTFR 292
 Db 237 HLYVDLAGSRQKRTAEGRLEKGININGLCLGIVISALDDDKGSPYRDSKLTFR 296
 QY 293 LQNSLGNKAKYIITCTIPV--SPDETISTQFASNAKHVRTPVNVETLDEALTKYR 350
 Db 297 LQDSLGSNHTLMKACVSPADSNLEETLSTLYADRAKTKKPTVN-----IDPHT 348
 QY 351 KEILDKQOLEN-----LESSETKAQAMAKE--EHTQLAEIKOLHREKREDRIWHLINI 403

Db 349 AELNHLKQYQVQVOLLQAHGGLPQSTNAEPSENLQSIEMKNQSLVEENKLSRC--L 406
 QY 404 VVASSQESQODQVKKRRRYTAPGKIQNSLHAGSVSDPMISRLPGNFSKRAKFSQMS 463
 Db 407 SKAQQTAQMLERIIITTEQVNEKLNKLEIRQHAACKIDL-----QKIVETLEDE 458
 QY 464 FPEIDSVV-----TEFSPFDALSMDSNGIDAEAMNLASKYTHREKS-----507
 Db 459 LKENVELICHLQOLITQLSDEYACT---AAADITVAEEBAQVETSPETSRSSDAFTQH 515
 QY 508 -LHOSMI--DEQGISDVQFHD-----SKENQUL--QYLPDQGDMAECHKASPEKEI 555
 Db 516 ALHQAOQSKREVVELNNALAKFALVRKQTNQNDQPIQPIQYQDNKIKNLEVEYNQKQK 575
 QY 556 TSLQOQLOKRE-----EKKELVQSEFKIALAEQQLSYKAKNLEMTNSRBSHIN 606
 Db 576 EELVRELQTKANNAQAKLSEHRKLLQLEQIDAKRLKLEQSKITL 625
 QY 607 AEVQTVKEKVEKEMSVLGDGYNASNDLQDSVDGRKLSHDECIEHRKMLEQKIV 666
 Db 626 -----KESTERTYSKLNQ-----EIMAKNORY- 648
 QY 667 DLEEFLENKKSSENDQKSSQEDFMESIQLCEALMAERANALELALMRDNFDNILEN 726
 Db 649 ---QLMRQKKEAEKFRQKQKRD--KEYIQKE---RDRKROYELKTLER--NQK---QS 697
 QY 727 EYTKR---EIAIDERSLKEN--QETNEFELLEKEQKE---HEAOLIHIGSLKLVENA 778
 Db 698 NVLRKTEEAANAANKRLKDALQKREVADKREKETSQSGNEGTAAAVRMNLGNEIEMVST 757
 QY 779 EYMNQNLDEEDLETKRLKEQETQLAELKRADNLQKVR--NFDLSVMG-----827
 Db 758 EBAKRLNDLLEDR--KILAQDVQQLKEKESRENPPIKRTCTFSLVGHVGLSEDCI 816
 QY 828 -----DSKELCEIYQLQKQSLSDAFAVTRDAKQEC-----SFLRSEN 864
 Db 817 TQOISLETEMELRSQAIDLQOKLIDAE--SEDPKQCMENIATILKAKLKLIGEL 874
 QY 865 LEIKEMEDTSNMWYQKEKAASLFEKOLETERKSNYKKNADQOKEL-----QSAFMEIYVL 920
 Db 875 VASKIHVTKLENSLQSKASCADQKMLFEQDNHSEITELQALVYMEQDHQEKVYL- 933
 QY 921 NGLLAGVPRDILLSVELLEKRVSEFSKOLEKALEKNALENEVTLSSBYKFLPNEVECI 980
 Db 934 -----LVSOLO--ESQMAE--KOLEKSASEKE-----POLVSTIQ 964
 QY 981 NQISKASEEIMLLKQEGESASIIISKQELIMQDQSEQLQTLDEYTHQSKVQOTEE---1037
 Db 965 CO-----DELEKARKVECEQNOQLOENETIKOKLI--LLQVASROKHLNPDTLLSPDSSF 1018
 QY 1038 QYLEKKKMHDDLEKRYIRNKSEADL-----LREMEMLKGTMESVE 1078
 Db 1019 EYIPKPKPSRYKKEFLQSDMDIEDLKQCSHVSVNEHEDGQGDGDSDE 1066

RESULT 8
 US-09-592-054-2
 ; Sequence 2, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684el. motor proteins and methods for
 ; TITLE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592,054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1232

```

: TYPE: PRT
: ORGANISM: Human
US-09-592-054-2

Query Match
Best Local Similarity 24.6%; Score 945.5; DB 4; Length 1232;
Matches 372; Conservative 275; Mismatches 489; Indels 369; Gaps 61;

QY 3 EGDAAVAVRVRPLIORE--GGDQANLQWKAGNNTISOVDGT--KSFNDRVFNHSTESQ 59
DB 6 KGIIVRALRCPPLPKPEISGCMCLSFVDEGPQV--VVGTDKSFYDFEVDSTEEOE 63
QY 60 IYDIAVPIIRSAOGNGTFAVGOTSSGKTVMG-----TPNSLIGTPOALDQVF 112
DB 64 VENTAVAPLKGVRKGNATVLAIGTSGKTYSMGAYTAEGENEPVGVIPRIQTLF 123
QY 113 KIIOEINREFILVSYMEIYNETVKDILCDRRKKPLEIREDEPNRYVVDLTELVMV 172
DB 124 KEIDKSDPEFTLVSYLEIYNEEILDLCPSRKAQININEDPKKGIKIVIGTEKTYLV 183
QY 173 PEHVIOIKKGEKRRHGETKANDHSRSHRTIFRMIVESRDRDPTNSGCGAVMSHL 232
DB 184 ALDVSCLEOGNNSRTVASTAMNSOSSRSRAILTISLEOGKKSCKNS-----FRSKL 236
QY 233 NLVDLASERASOTGAGVRLKEGCNINRSLFIIIGVYTKKISDQAGGFIVYRDSKILRI 292
DB 237 HVDLASSEROKTKARDBLKEGININRGLCLGNVYISALGDDKGGFADYRDSKILRL 296
QY 293 LONSLGNAKTVIICITPV--SEDETSLTQFASAKHVNTPHVNEVLDEALKKRYR 350
DB 297 LQDSLGNSHTMTACVSPADSNLEPTLNTLRADRAKIKNPV--IDPQT----- 348
QY 351 KEIIDLKOLEENLESSESTKAQAKKEBHOLLAEIKOLKEREEDRLWHLTNIVVASSQE 410
DB 349 AELNHLKQVOYL-----QVLL----- 365
QY 411 SQQORVRRKRYTWAPGKIQNSLHASGVSDFDMLSRPGNFSKAKPSPSPPEIDDS 470
DB 366 -----LQAHG-----GTLPGSITE-----PS----- 382
QY 471 VCTEFSDFDALSMDSNGDAEMNLASKYTHREKTS--LHOSMIDFGQISDSVOFHDSS 528
DB 383 -----ENLOSIMEKN-----QSLVEENKLSNGSEAGGQTAQMERIIMTQQA 426
QY 529 KENLOLQLPDSDGMACRCR--ASFEKETISLQOOLQSKSE--BEKEIYVOSFELKAELEQ 586
DB 427 NE-----KKNAKLELRHAAKLDLQKLVETLEDEKENEVEIICNLQOLLTQJLDE 479
QY 587 LSVAKMLMEWTNSREHSINAEVOTDVEKEVVRKREMSVIGDSGYNASNSDLQDSSVDQKR 646
DB 480 -----TVACMAAAIDVAVDEAOVETSP-----TSRSSDAFTTQHALR 518
QY 647 LSSSHDECIEHRKMLEOKIYDLLEEFIEINLNRK--SENDKQ--KSSBODEMESIOLCEAIAAE 704
DB 519 QAQSKRELYELNKLALK-----FALARKMTQNDSQLPIQYOYQDNT----- 561
QY 705 KANALELALMRDNFDNIILENETLKEIADLESKLENQNTNEFEIIEKTOKEHEHQ 764
DB 562 KEPELEVINIOLQEK--EELVYELQTKKADANQAKLSERRRKLQLEEQIADLKKR----- 615
QY 765 IHEIGSKLIVENAMETNONLEEDLETKTKLBOEIOALRLKRRADLQKRVNFDLSV 824
DB 616 LNEOSKILLKESTERYVSKINOEL-----RMMKNQRYOL--MRQKKDEAE--KFFQW----- 664
QY 825 SMGSEKICEIEIFOLKOSLSDAEAVTRDAQKESFLSENLELKEKEMEDTSNMYNOKEKA 884
DB 665 -----KQKRDREVIOLKE-----HDKRRQ-----YELLKLEBNPQKSNVLRKKEE 706
QY 885 ASLEKOLETEKSYKKMEADLOKELQSAFNEIYINGL--LAGKVPBLLSRVLELEKY 942
DB 707 AAANKRILKDALQKOREV--ADKRRKETOS-----RGMESTAARVKMWMLGNEIEVWYST 757
QY 943 SEFAKOLEKALEKNALENETVCLSEYKFLPNEVECLKNQJSKASEEIMLKQSGEHSAS 1002

```

```

DB 758 EBAKRLNDLLEDRKILQDVAOLKE-----KKEGGENPP 793
QY 1003 IISKOELIMOESEQIOLITDEVTHTQSGVOOTEBOYLEMKRMHDLFEKYIRNKSEAD 1062
DB 794 KLRRTEFLTEVRGVSSEDSIT--KQLESLETEMEFPAQIA--DLQOKLL--DAESED 848
QY 1063 LLRE-----MENIKGMESEYKVIADTKHELE--TIDKQQLHEKKYF 1105
DB 849 RPKQWENIATILEAKCALKYLIGELVSKIOVSKLSESLSKQTSKSCADQMKMLEERNH 908
QY 1106 FOAMOTIPIPIPLSDLPKSLVEGNSODPIEINDYHNLIALATERNNINVCLETENSL 1165
DB 909 FAEIFTEL-----QAEIVRMEQHPKV-----LYLLSOLQOSQMA-----EKQ 948
QY 1166 KEQVIDNTLOLOS--LOAGSIEKSDLOKRPDLEGEVRLLEMLLKLHLDLSOLSIEKL 1224
DB 949 EESVSEKQOQLOSTLKCQD---EELKMRVEQEOHQ--QLRBNETIKQKLTLLQVASRQK 1004
QY 1225 QLEN---LEVTEKQTLQEBKKNITIERN--ELQTNFEDLK-----AEHDSLKQDSE 1272
DB 1005 HLPKOTLLSPDSSEFYVQPKRPSRKKEKLEQSMIDIEDLKYSHEHSDGDDDE 1064
QY 1273 NIEOSIETODELRAAOEELR-----EOKOLVDSFROQLLDCSVG--ISSPNHDAVANO 1323
DB 1065 GDDEEMKPTKLVNYSRKNIQGCSCGKMGCKNQ--GCHKQKSDCGVDCCPPTKCRNRQ 1122
QY 1324 EKVSIGEVNLSQSEMLNGERDELQTSICALVSELELLRAHVSVGENLEITTKLNGLEK 1383
DB 1123 GKDSIGTVERNQ-----DSSESK-----LE-----DPTVYTPGLSFENP 1157
QY 1384 EILKSESEVYLSM--LENLKE-----DNKKLEQAEY-----SKENQFSL-- 1425
DB 1158 --VCATPSKILKEKMDVEQVLSKTPPAPSPFDLPLELKHVATEYQENKSGKKRRALA 1215
QY 1426 --EEVFSGSKVLDE 1438
DB 1216 SNTSFFSGCSPIEER 1230

RESULT 9
US-09-722-139-2
: Sequence 2, Application US/09722139
: Patent No. 635471
: GENERAL INFORMATION:
: APPLICANT: Beraud, Christophe
: TITLE OF INVENTION: No. 6355471el motor proteins and methods for
: TITLE OF INVENTION: their use
: FILE REFERENCE: 1055
: CURRENT APPLICATION NUMBER: US/09/722,139
: CURRENT FILING DATE: 2000-11-24
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 1375
: TYPE: PRT
: ORGANISM: Human
US-09-722-139-2

Query Match
Best Local Similarity 25.5%; Score 906; DB 4; Length 1375;
Matches 413; Conservative 247; Mismatches 531; Indels 430; Gaps 71;

QY 6 AVKVCVRVRPLIOREGDQAN--LQWKAGNNTISOV-----DGTKSFNFDRVF 51
DB 3 SVKVAVVRVRPNRRREKDLKAFIIQMEKSKTITMLKIPGEGTQDSGSRKRTFTYDPSF 62
QY 52 NSHSTS-----QIYDIAVPIIRSALOGNGTFAVGOTSSGKTIVTMGTPNSLGI 104
DB 63 YSADTKSPDYVSOAEVWFETLCTDVVYKSAFEGYNACVFAYGOTGSGSKSYTMGNSGSGSLI 122
QY 105 PQALDQVFKIIQEI--PNREFILVSYMEIYNETVKDILCDRRKKP-----LEIREDFN 157

```

```

Db 123 PRICGLFSRINETRWDEASPRTEVSYLEITNERVRL--RRKSKTFLNRREHPK 179
QY 158 RRVYVADLLEELVMPBENVYQWIKKGEKNRHYGETKANDHSRSTTF--RMVYESDRN 215
Db 180 BGPVEDLSKHLVQYGVVEELMDAGNINRTTAAGMNDVSRSHAITFTIKFTQAKFDE 239
QY 216 DPTNSDCGAVMVSHLNVLDLAGSERASOTGAEGVRLKEGNCINRSFLITGOVIKKLSD 275
Db 240 MP-----CE--TVSKHLVLDLAGSERADNATGATVRLKEGNCINRSFLITGOVIKKLSD 289
QY 276 GOAGGFYIRDSKLTIRIIONSLGNAKVIITCTTP--VSFDELSTLQFASAKHVRNT 333
Db 290 -----FVYERDSVLTWLLKDSIGNSKTIATISPADVYNGETLSTLRVANKNITNK 344
QY 334 PHVNEVLDEALLKRYKEIIDLK-----KOLENLESSSETKQAMAKEHTOL-LAEI 386
Db 345 PFINDAVAK-LIRELRAEIRAKLTILAOGNQIALDLSPT-----ALSMERKLOONEARV 398
QY 387 KOLHKEKREDRIWHLNIVAVASSQESQDQRYKRRRYTMAPGKIONSLHAGSVDFMLIS 446
Db 399 QELTFEMTNK-WNET-----QNTLKEQTL-----ALRKEGIG----- 429
QY 447 RLPGFSSKAKFSDMPSPPEIDDSVCT-----ERSDDDLASMD--SNGIDA 492
Db 430 -----VIDSELPHLIGIDDLSTGILYLHLKREGQTVGHDASTEDQIVLHGLDL 481
QY 493 E-----WNLASKYTHREKTSIHQSMIDFQISDVQFHDSS-----KENOLOY-LPRDS 540
Db 482 ESEHCIFENIGCVT-LIPISGQCSVNGQIVATHLNGQAVILLGRTNFRNHRKEA 540
QY 541 GDMACR-----ASFEKEITSLQOOLOSKEFEKELYQ-----SFEIKIAIEEQLSVK 590
Db 541 AKLREKRSGLISSFSLSMDL-----SKSRNLSAVMLXNPGLEFEROORELEKLESK 595
QY 591 AKNLEMYNRSRHSINAEOQ-TDYKEVVRKEMSVLGDGYNASNDLODSSVDGRKLSS 649
Db 596 KRLIEM-EKQKSDKALEMROQEVETQKRETIY-----OLO-----IR 635
QY 650 SHDCIHRKA-LQKIVDLBEETENLNKSSSENDKQSSQODEMESIQCEAIAEAKANA 708
Db 636 KOESLKRSHFNKTKDL-----LAKEKEFEERLRQ----- 670
QY 709 LEELALARDNDNITLENLTKREIADLER-SLKENQENNEPEILEK-ETQKEHEAOL 764
Db 671 -QELTELQKR-----QEEFLPRLQOEBLOLRIKELNNKEKEKFOIQLDOLKEKDEOY 724
QY 765 IHEIGSLKLVENAE--MYNOLLEEDLETK--TKLKEOEIOLAEIRKRDNLQKV 817
Db 725 A-KLELEKRRLEBOEKQVMLVAHLEBQLRKEQEMQLLRGEVQWEEERK----- 775
QY 818 RNFDSLVSMDSEKICEIIFOLKQSLSS--DAVATYRQAKQCSFLRSNLEKEMED 873
Db 776 -----DLGIRISLLIRVEKARAGGDEDELEKALRFEFEKRRQIVLKLVLNED 825
QY 874 TSNWYNOKERASLFEKOLETEKSNYKMEADLOKELASFNFINVLNGLACKVRDIL 933
Db 826 LVQOKDLKEVQEOEILFCLNCEHDKESRLLEKHDSTVD-----VTEVQODE 876
QY 934 SRVLEKYSSEKOLEKALEENALENEVTCLSE-----YKFLPNEVDELKN--QISK 985
Db 877 KIRPVERILOQYKEROLO-----YLLONHLPTLLEEKORAFELIDGSLDITLQYOKER 930
QY 986 ASEEMILKQEGHSASISIKOELIMOEOSE--QILOLTDVY--HTQSKVOQTEE 1037
Db 931 EME-----KEQOLQOYQANANQLOKLTATEFTFNINARQOEKVKRKKK 974
QY 1038 QYLEMKRMHDLFEKYIRNKSSEADLLIREMENLKGTESEVYIADTKHE-LEETIROE 1096
Db 975 ELLES-----REKQOREALERALARLRRHSLQ-----RSTLGTETEEOER 1016
QY 1097 QLLHEKKYFFQAMQTIFFITPLSDSLPFSKLYEGNSODPLEINDYHNLIALATERNNIV 1156

```

```

Db 1017 Q-----KIASLNSG-----SREQSGLO 1034
QY 1157 CLETERNSLKQVIDLNTLOSLQASIERKSDLOKPRQODEEGEV-KLLEMEILLKHTL 1215
Db 1035 SLEAEQALEKQDQRLREYELQLOKQIYEVQVQKDHGHLLEGVAVASSSLPYAENSHLV 1094
QY 1216 ---DSQS--TEKQLENLEYTELQTL-----QEMKNITIERNELO 1253
Db 1095 PLAMARINAYIEE-----EVQRLQDLHARVIEGCGTSADTKMDKELNGTQRLKY 1148
QY 1254 TNEFDLK-----AEHSLKODLSENIEOSIETODEL-----RAQOEELREOK 1295
Db 1149 ELCDLDCVLMPEPDAAACNHLLODL--VOLSDMKTEIPLDVLPRGVYSSKFQ 1205
QY 1296 QLYDSFQOOLLDGCVGISSPNHDVANOQEVSLGEVNSIQSEMURGERDELQTSKALVS 1355
Db 1206 TLVDMI--YFLHGMENVDP-----SLAEVOLLTYTVKVGDSGHQCCQSLV- 1251
QY 1356 ELETTRAHVSVE-----GENLEITKRLNGLEK-----EILKSESESY 1394
Db 1252 ---LNLTHIALVYEDCYFTYRISRNIPPGAOFDVI-KCHALSERCVYVPEKKNVSTV 1307
QY 1395 LKSMLEMLKEDNNKLEQAEYSSKEN--QFSLEEVFSQSKLYDEIYLAQKAAER 1452
Db 1308 ELVFLQKLPKPSVGSRSNPPHELQEAPNVOLFTPLVLYQSQVAP--EYWKLTFFNSQDA 1365
QY 1453 L 1453
Db 1366 L 1366

```

```

RESULT 10
US-09-721-832-2
; Sequence 2, Application US/09721832
; Patent No. 639346
; GENERAL INFORMATION:
; APPLICANT: Berand, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: NO. 639346el motor proteins and methods for
; FILE REFERENCE: 1055
; CURRENT APPLICATION NUMBER: US/09/721,832
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: Human
US-09-721-832-2

```

Query Match 6.1%; Score 906; DB 4; Length 1375;
 Best Local Similarity 25.5%; Pred. No. 6,7e-37;
 Matches 413; Conservative 247; Mismatches 531; Indels 430; Gaps 71;

```

QY 6 AKYCVRVAPLQREGDOAN--LQWKAGNNTISQV-----DGTSENFNRV 51
Db 3 SVKAVARVAPMNRKRDLEAKFLIQEMSKYTIIMLAKIPEGGTGDSGRERTFTYDSEF 62
QY 52 NSHESTS-----QIYOEIAVPIIRLSALOGYNGTIFAYGOTSSGKTYMTMGTSILII 104
Db 63 YSADIKSPDYVQENAVFTLIGDVYKSAFEGYNACVFAYGOTSGSKTYMTMGNSGDSGLI 122
QY 105 POAIOEVEKIIQEI--PNREFLRVSYMEIYNETVKDILCDDRKKRP-----LEIREDFN 157
Db 123 PRICGLFSRINETRWDEASPRTEVSYLEITNERVRL--RRKSKTFLNRREHPK 179
QY 158 RRVYVADLLEELVMPBENVYQWIKKGEKNRHYGETKANDHSRSTTF--RMVYESDRN 215
Db 180 BGPVEDLSKHLVQYGVVEELMDAGNINRTTAAGMNDVSRSHAITFTIKFTQAKFDE 239
QY 216 DPTNSDCGAVMVSHLNVLDLAGSERASOTGAEGVRLKEGNCINRSFLITGOVIKKLSD 275

```



```

Db 240 MP-----CE---TVSKIHVLVDLAGSERADATGATGVRLEKGNINKSLVTLGAKKKQV-- 289
Qy 276 GAAGGFINRDSKLTIRILNLSGNNAKTYIICTTP--VSPDETSTLOFASTAKHVRT 333
Db 290 -----FVPRDSVLWMLKDSLGNSKTMITATISPADNYGETISTLRYANRAKNIINIK 344
Qy 334 PHVNEVLDEALKRYRKETIDLK-----KOLENSESSEKKAQMAKEHTOL-LAEI 386
Db 345 PINEDANVK-LIRELRAIARLKTLLAOGNOIALDSPT-----ALSMEKLOONEARV 398
Qy 387 KOLHEREDRIWHLNTIVASSQSOODQVRKRRVTAPEKIONSILASGVSDMLS 446
Db 399 OELTKEWTK-WNET-----ONILKEQTL-----ALRKEIG----- 429
Qy 447 RLPGNFKAKRFSMPSEFIDDSVCT-----EFSDFDALSMMD--SNGIDA 492
Db 430 -----VLDSELPRLHIGIDDLSTGIIYHLEKQTYVGRDASTEQDILVHGLDL 481
Qy 493 E-----NMLASKVTHREKTSLSHOSMIDFGQISDSVQFHDSS-----KENOLQY-LPKDS 540
Db 482 ESEHCLFENIGTVT-LIFLSSQCSVNGVQYVEATHLNOGAVIILGRIMRFRNPKRA 540
Qy 541 GMAECRK---ASPEKETISLQOOLQSKKEEKELVQ-----SEELKIALEBOLSVK 590
Db 541 AKLREKRKGLSSPSSLSMTDL-----SKSRENLSAVMLYNGLEPERQORELEKLESK 595
Qy 591 AKNLEKVTNRSHSTNAEQ-TDVEKEVYRKEMSVLGDSGYANSDLODSSVDGRKLS 649
Db 596 RLILEEM-BEKQSKAELERMQOEVEYQKETEYV-----QLO-----IR 635
Qy 650 SHDECIEHRKM-LEOKIVYLEFIEENLKNKSENDOKSSEODFMESIQCEAIMAKANA 708
Db 636 KQEBESLKRSFHTENKLDL-----LAKEKFEERLEBQ----- 670
Qy 709 LEEALMRDNFDNIILENTLKRITADLER--SLKENQTFNEFELK--ETQKEHAOL 764
Db 671 -OEIELOKKR-----QEEETFLVQOELORELKLNNEKAEKFOEIOEDLOKDEQY 724
Qy 765 IHEISLKLIVENAE---MYNONLEEDLETK---TKILKBEIQOAELEKRAKADNOKKV 817
Db 725 A-KLELEKRLKEOEKQOYVAHLEBQUREKQEMIQLRKEGVQVVEEKR----- 775
Qy 818 RNFDLSVSGDSEKCEETFOLOKSTLS---DAEAVTRDAQECSEFSENELKEMKD 873
Db 776 -----DLEGIRRESLRYKVARAGDDEGELEKQALFFEFKRRQYKVLVLEMD 825
Qy 874 TSNMTNQEKAASLEKOLETEKSNYKMEADLOKELQSAFNEINYLNGLAGVPRDIL 933
Db 826 LVQOKDILKKEVQOEQOELIECLCEHDKESRLLEKHDESVTD-----VTEVQDRE 876
Qy 934 SVELEKQVSESKOLEKALEKKNLENVTLSE-----YKFLPNEVECLKN--QISK 985
Db 877 KIKPVEYRQYKERQLO-----YLLQNLPTLLEKQARAFILDRGPLSLDNLTYQVER 930
Qy 986 ASEELMLKQEGEHSASISKOELIMQOSE---QILOQLEBVT---HTOSKVOQTEE 1037
Db 931 EHEE-----KEEQLOAQYANANOLOKLOATPEFANINAROEBEKYKPKKK 974
Qy 1038 QYLEKKNHMDLFEKYIRNKSAEDILREMNILKGTMESVEYKIDATKHE-LEETIRDE 1096
Db 975 EILLES-----REKQOQREALRALRLERHSHALO-----RHSTGLTEIEEOR 1016
Qy 1097 OLLHKKYFFQAMQTIPTPLSDSLPSKLYEGNSQDIEIENDYHNLALATERNNIIV 1156
Db 1017 Q-----KLASLNSG-----SREOSGLQA 1034
Qy 1157 CLETERNSLKEQVINDLNTQLOSLQASLEKSDLOKPRKODLEGEV-KLLEMLLKHILT 1215
Db 1035 SLEAGQEALEKQOERLEVEIQOLKQYEVQVQKHGHTLGGKVAASSLPAESAESHV 1094
Qy 1216 ---DSQLS--IEKLOENLEVTETKLOTL-----QEEKKNTITIERNEIQ 1253
Db 1095 PLMDARINAYIEE-----EYORRIODLHRVISEGQSTADTMKDNKELHNGTIOBKLY 1148

```

```

Qy 1254 TNFEDLK-----AEHDSLKODISENIEOSIETODEL-----RAAOELRECK 1295
Db 1149 ELCRDILLCYLMBEPDPAACANHLLQODL---VQSLDKMTIPLDVLNGVQVSSKFT 1205
Qy 1296 QLVDSFRQOOLDCSVGISPPNDHADVANOKEVSGEYNSIQSEMLNGERDELOTSCALVS 1355
Db 1206 TLVDMT--YFLNGMNEVNP-----SLAEVQOLLTYTVYMGDSGHDQOSLV- 1251
Qy 1356 EELLRHAKVSYE-----GENLEITKRLNGLEK---ELIGKSESEV 1394
Db 1252 ---LMTHTHALYKEDCFYPRIRSRNI PPPGAQFDVI-KCHALSEFRGVVPEKKVSTV 1307
Qy 1395 LKSMLENTKDNKKLEQAEYSKREN--QFSLEEVSGSKLYDEIVLKQOLKAABER 1452
Db 1308 ELVFLQKLPVSGSRNSPPEHLOEADNPVOLFPTPLYLQSGONVAP--EVMKLTFSNDBA 1365
Qy 1453 L 1453
Db 1366 L 1366

RESULT 11
US-09-721-689-2
; Sequence 2, Application US/09721689
; Patent No. 6440685
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6440685el motor proteins and methods for
; FILE REFERENCE: 1055
; CURRENT APPLICATION NUMBER: US/09/721,689
; CURRENT FILING DATE: 2000-11-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1375
; TYPE: prt
; ORGANISM: Human
US-09-721-689-2

Query Match 6.1%; Score 906; DB 4; Length 1375;
Best Local Similarity 25.5%; Pred. No. 6.7e-37;
Matches 413; Conservative 247; Mismatches 531; Indels 430; Gaps 71;

Qy 6 AVKCVVRVRLPIQREGQDQAN--LQWKAGNNTISQV-----DGTSSFNDRVF 51
Db 3 SVKVAVRVPRMNRREKDLKAKFLIOMEKSKTITNLKIPGGTGDSGRERTKFTYDFSG 62
Qy 52 NSHESTS-----QIYQEIAPPIIRSLAQYNGTIFAFGOTSSGKTYTMAGTPNSLGI 104
Db 63 YSADTKSPDYVSQEMVFKLTGIDVYKSAFEGYNACVFAFGQSGSGSYTMGNSGDSGLI 122
Qy 105 POAIOEVEKTIQEI--PREFLLRYSYMEIYNETYKDLICDORRKP---LEIREDEN 157
Db 123 PRICGELFRINETTRMDASFRTEVSYLEITNERKRDLL---RRKSSTFMLRPREHK 179
Qy 158 RNVYVADLTELIVWPEVHIQWIKKGEKNRHXGETKMDHSSRSHTF--RMIVSRDN 215
Db 180 EGPYVEDLSKHLVQYNGVDEELMDAGNINRTAATGMNDVSSRSHAIFIKFTQAKFDS 239
Qy 216 DPTNSNCGCAWWSHLNLVYDLASGRASQTAQEVRLKEGNINRSLFTILQOVIKKLSL 275
Db 240 MP-----CE---TVSKIHVLVDLAGSERADATGATGVRLEKGNINKSLVTLGAKKKQV-- 289
Qy 276 GAAGGFINRDSKLTIRILNLSGNNAKTYIICTTP--VSPDETSTLOFASTAKHVRT 333
Db 290 -----FVPRDSVLWMLKDSLGNSKTMITATISPADNYGETISTLRYANRAKNIINIK 344
Qy 334 PHVNEVLDEALKRYRKETIDLK-----KOLENSESSEKKAQMAKEHTOL-LAEI 386
Db 345 PINEDANVK-LIRELRAIARLKTLLAOGNOIALDSPT-----ALSMEKLOONEARV 398

```

```

QY 387 KOLHREDRHWTNIVVASSQSOQDQRYKRRVYAPKIQNSLHSGVSDPMLS 446
DB 399 QELTKEWINK-WNET-----QNLKEQTL-----ALRKEGIG----- 429
QY 447 RLPGNFSKAKFSDMPSPEDDSVCT-----EESFDALSMMD--SNGIDA 492
DB 430 -----VVLDSLPRLHIGIDDLSTGIIYLHLKEGQYVGRDASTEDQIYLHGLDL 481
QY 493 E-----WNLSKYTHREKTSIHQSMIDGQISDSVOFHDS--KENOLQY-LPND 540
DB 482 ESEHCFIFENIGTVT-LPLSGSQSVNGVOIVEATHLNGQAVIILGRTNMFRENHPEKA 540
QY 541 GDMACRCR--ASEFEKITSLOOQLOSKREBEKKELVQ-----SFEKIAELEBOLSVK 590
DB 541 AKLREKRSGILLSPSLMTDL-----SKSRENLSAVMLYNGLEFEHQBELEKLESK 595
QY 591 AKNLEMYNSRHSINAIVQ-TDVEKEYVRKEMSVLGDGYNASNSDLODSSVDGRLLSS 649
DB 596 RKLIEEM-EEROKSDKALEERMOQEVETORKELEIV-----OLQ-----IR 635
QY 650 SHDECIEHRKM-LBQIYDLBEEFLENLKNKSENDKOSSEODFMESTQLCAIMAKRANA 708
DB 636 KOESLKRSHFHEIKLKL-----LAKKEFEERLREO----- 670
QY 709 LEBLALMDFNDIILENETLKEKRIADIER--SIKENQETNEFELEK--ETOREHEAOL 764
DB 671 -OEIETLOKRR-----QEETFLRYOEELORLKEILNNKAKERFQFOLDLOKREKBOY 724
QY 765 IHEISLKLLENVNAE---MYNONLEEDLETK---TKLKQELQALMERKRAANLOKRY 817
DB 725 A-KLELEKRLLEBQEKQEVYVAHLEBOLREKQEMTOLLRGEVQWVEEER----- 775
QY 818 RNFDLSVSMGDSKLCBEEIFOLKQSLS---DABAVTRDAQKESFLESENLELEKEMED 873
DB 776 -----DLEGIRRESILAVKKEARAGGEDGELEKAQURFEFRKQRYLVKLVNEKD 825
QY 874 TSNMYNOKERKASLEKQLEFEKSNYKMKEMADLOKELOSAPNETIYNGLAGVPPDOL 933
DB 826 LVQOKDILKKEVQOEQELTECLCEHKEESRLKEKHDSVYD-----VTEVPQDFE 876
QY 934 SRVELKKEVSEFSKOLEKALEEKNALENEVYCLSE---YKFLPNEVECKLN---OLSK 985
DB 877 KIRVEYRLOKQKEROLO-----YLLONHLPTELEKORAFEILDRPDISLDTLYOYER 930
QY 986 ASEIMILKOBEGHSASIIKQETIMOQOSE---QIILQTLDEVT---HTOSKVOQTEE 1037
DB 931 EMEB-----KEEQLOAQYQANANOLKLOQATFEFTANIMARQOEKVRKKEK 974
QY 1038 QYLEMKKMHDDLPKRYIRKSEADLLREMNELKGTMESEVYKTIADTKHE--LEETINDKE 1096
DB 975 EILES-----REKQORALERALARLERHRSALO-----RHSITLGTIEEOR 1016
QY 1097 QLEHKKYFFQAMQITPITPLSDLPSPKLVESQDPIEINDYHNLMLATEPNNIMV 1156
DB 1017 Q-----KLASLNSG-----SREQSIOQA 1034
QY 1157 CLETERNSLKEOYDLTFOLOSLQASQIESDLOKPRKODLEGEV--KLLEMEILLKGHLT 1215
DB 1035 SLAEQORALEKDERLEFEYEQLOKQIYEVGYQKDHGHTLEGVASSSPVSAKESHLY 1094
QY 1216 ---DSOLS--TEKLOLENLVEYTELOTL-----QEMENITIERELO 1253
DB 1095 PLMDAKINAYIEE-----EVORLODLARHVISGEGSTADTKMDKNKLNGTIOKRLKY 1148
QY 1254 TNEFDLTK-----AEHDSLKODLSLENIEOSIETODEL-----RAAQEELREOK 1295
DB 1149 ELORDILCYLMPERDAACANHPLOODL---VOLSLDMKTEITPDLVLPNGVYSSKFOJ 1205
QY 1296 QLVDSFRQOLLDSCSVGISSNHDAVANOKEKVSLSGEVNSLOSEMNGERBELQTSKALVS 1355
DB 1206 TLVMDI--YFLHGNMEVNV-----SLAEVOLLILTYTVKMGDSGHQDCSIV- 1251

```

```

QY 1356 ELEILBAHVSVE-----GENLEITKINGLER-----EILKSESSEV 1394
DB 1252 ---LUNTHIALVKEDCYFYPIRISRNIPPGAOFDVI-KCHALSEFRCVVPEKKNVSTV 1307
QY 1395 LKSMLENLKEDNNKIKQEAQESYSSKEN--OFSLSEVFSQKLVDETEVLKQALKAER 1452
DB 1308 ELVELQKIKRPSVGRSNRSPPEHLQAPVQLETTPLYLQSGQNVAP---EYVKLTFFNSODEA 1365
QY 1453 L 1453
DB 1366 L 1366

```

```

RESULT 12
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

```

```

Query Match 5.98; Score 866; DB 4; Length 10182;
Best Local Similarity 18.6%; Pred. No. 6,3e-34;
Matches 661; Conservative 697; Mismatches 1295; Indels 898; Gaps 157;

```

```

QY 15 PLIOREBQDQANLOWK-----AGNN-----TTSOVNGTSPFNDRY-----FNH 54
DB 5623 PINKNEIEBQAINQINTQYALSGENKLRHDOESTNRQLEGSLTAQIMAEKDLVNOA 5682
QY 55 ESTSQIYOET-AVPIIRSALOGNGTTFAYGQTSQKRYTMGTGPSLGIIPAIOEVER 113
DB 5683 KTRTDVAQKLAKEINSASNLRDGIQNKEDIKSSAV--INADPKRYAYQALQALMAN 5741
QY 114 IIOEIPRER-----LIRV-----SYMEIYN-----EYKDL-LCDDRKKPL 150
DB 5742 TINATPVELNKATTFQALSRYQAOQODLDGYOQALNARQOATQTVNGINSILNDGOKREL 5801
QY 151 -----EIREDFNRNVYADLTLELVMPBEHVIQ-----WIKKGENRHYGETK 193
DB 5802 NLLINSANTRTKQOEELNKATELNHAMEALRNSVQVADVQKSSNYVNEQEPQHNYDNA 5861
QY 194 MND-----HSSRSHITFRMIVESRDRNDPTNSENCDGAVVSHNLVLDLAGESASQTA 248
DB 5862 VNEAQATINNNAOPVLDKIAIERLTQTVMYTKDALGCA-----QKLTQOQAAETGI 5913
QY 249 EGYR-----LKEGINSLSPLTGOYIKRLSDGQ-AGGF 281
DB 5914 RGLTSLNEPQKNAEVAKVTAATTRDEVRNIRQEAATLDTAMGLRKSISDKNDTKNSRY 5973
QY 282 INTRDSKLTRILONSIGNAKTYIICTIPVSPDETLSTL-----QFASAKHVRNPH 335
DB 5974 IN-EDHDQOQADNAV-ANAOYI-----DEQATLSDDTINQLANAVTOAKSNLH 6022
QY 336 VNEVLDEALLRKRYKEIILDKQLENIESSESSEYKQAQAKKEHRO----- 381
DB 6023 GPTKLOHD---KSAKQTL---AQLONSAQKHMEDSLIDNESTRTOQHDLTEAQLD 6076
QY 382 -LLAEIKOLHKEREDRIWHLTNIVV-----ASSQESQO-----DQRYKR 419

```

Db 6077 GLMGLKESIKD-----YTINIVSNGNYINAPESKROAYDAVQAONNIINGTNOPTIN 6129
 Qy 420 KRRTWAPGKIQNSLHA-----SGVSPFDMLSRLPGFSSKAK---FSDMSEF 464
 Db 6130 KGNVTTAQTQVKNTRKDALDGDHRLLEAKNMANQTRNLISNL--NNAQKAEMKLVNASTL 6188
 Qy 465 PEIDDSVCTERSDPDALS-----MMDNSCIDAEWNLASVYT 501
 Db 6189 EOVQONLOT-AQOLDNMGELRQSIACKDYKASKYINEDPOIKQNTDYDAVORVETIIN 6247
 Qy 502 HREKTSIHSOMIDFGOISDSVQ-----FHDSKQENQ-LQYLPKDSGDMACRKAFFEKEI 555
 Db 6248 ETQPELLEKANID--QATQSVQNAEQALHGAEKLNQDKQTSSTELDGLTDLDAQREK- 6303
 Qy 556 TSLQOOL---QSKEEKEKELVQFELKA--ELEBQLSVK-----590
 Db 6304 --LREQINTNSRSDIKOKIKOAKALNDAMKLEQVQAKGVHANSPTYNEDSAQKDAY 6361
 Qy 591 ---AKNIE-AWTSNREHSINAEVOTDYEKEVVRKEMSVLGD-----SGVNA 632
 Db 6362 NNALKQEDLIINSSNPNLNAQDITLNLNLIKQADNLHGAKQLOQDKNTTNOAIGNLNH 6421
 Qy 633 SNSDLODSYVDGKRLLSSHDECIEHRKALEQKIVYLEEFENLNKKSSENDKQSSQEDFM 692
 Db 6422 LNOQOKALLQAIINGATSROQVAKKIKAEAA---LDEAMKQLEDQVQDDQOISSSPFI 6477
 Qy 693 ESIOLCERIAEKANALEBELALMRDNPDNITLENETL--KREIADLENSLKE-----742
 Db 6478 NEDSDKQKYNDKIOAKKEIINO-----TSNPTLDKOKIADTLONIKAVANNLHGDQ 6529
 Qy 743 -----NOETNEFELEKETQREHEAOIIEHIGSLKIVENAEWNOMLLEEDLETKT 793
 Db 6530 KLAOSKODANQNLHLDLLEE-QKNHFKPLINNAOTDEVENKOLEIAKQ-LNGDMSLH 6587
 Qy 794 KLAQEOTLAELRK--RADNLOKRVNPNFLSVSNG-----DSEKICEITFOLK 840
 Db 6588 KVINDKP-QIOHLSNYINADN--DKQVYDNAIKEAEDLIHHPDLDLHKALQDLINKID 6644
 Qy 841 OSLSDAEAVNR-----DAQECSFLRSENLEKEKEMDTSNWYNOKKASLEPEKOLE- 893
 Db 6645 QAHELNGESRFPKQALDNLALDIDSLNSLNPQOTVAKDNINHNHTTLESIAQELQAKEL 6704
 Qy 894 -----TEKSNYKMEADLOKELOSAFENEINY-----919
 Db 6705 NDAMKARHDSIMNOEJIRKNSNTYNEDLAQONAYNHANDKINHIIIGEDNATMDPIIITKA 6764
 Qy 920 -----LNGILAGVPRDLISRYELEKVSSEFS-----KOLEKALEKNALENEVTC 965
 Db 6765 TODINTAINGINGDQKLOD--AKTDAKQOITNFTGLTEPQKQALENIIINOQTSRANVAKQ 6822
 Qy 966 LSEKKEFLPNEVECKKNQISKASEIIMLLKQGEHSASITISQOELIMO--ESEQILOLTD 1023
 Db 6623 LSHAKFLNGKKEELKAVAAKS---LVRONSYINEDVSEKAEVEQOIAAGQELTINSEN 6878
 Qy 1024 EYVATQSKVOOTEQYLEMK--MHMDLFEKYIRNK--SEAEIDLK--EMENLKG-----1072
 Db 6679 NPTISSDINTIOEINDABENLHGD-----KLRQAQSIANKNEIOLNDGLNSAQITK 6931
 Qy 1073 -----TWESVEVKIADTK-----HELEETTRDKOEULLHEKKYF-----1105
 Db 6932 LIODIGRTTTPATQKLEEKAQINQAMQLOKSIOADKALINSSNYLNEDSEKKLAVDN 6991
 Qy 1106 --FQAMQTIPTPLSDSLPSKLVESQDPIELND-YHMLILATPR--NNIMVCIET 1160
 Db 6992 AVSQAEOI--INQLNPDYMDISNIQAITQKVIQAKDSLHANKLAQNAQDSNLITINOST 7048
 Qy 1161 ERNSLKEEVID-----LMTQLOSLQOASIEKSDL-QKPRQDLEE 1198
 Db 7049 NUNOKQALNDLNLNHQTKQOVAEIIIAQANKLNNEKQTLTIVEGQSNVQOQSKYINED 7108
 Qy 1199 GEVALLLEMLLKGH-----LTSQLSLEKTL--OLENLEVT-----EKLQTLQOEMKN 1244
 Db 7109 POYONIVNDISIQGREILNGTTDDVLNNKRTADAIGNIHLTKNDLHGDKLOKKAQODATN 7168

Qy 1245 ITIERNEL-QTNFEDLKAENHSL-----KODISENIBOSIETQDELRAPQEL---REOK 1295
 Db 7169 ---ELNVLTLNNSQROSEHEDINSAPRTEVSNLHNKAKLNEAMKOLEVALENSVK 7225
 Qy 1296 QLVDSFR-----QOILDCSVGISSPNDAAVANOEKVSLGCVNSIQSEMLRGE 1342
 Db 7226 KLSDFINDEDAQONETYSNALQKAKDIINGVPSITDKATIED--ALLEQONAR-ESLHGE 7282
 Qy 1343 RDELQTSKALVSELELLRAHVKSVEGENLEITTKLNG---LEKELIGKSESEVELKSM 1398
 Db 7283 Q-KLOEAKQNAVAEIDNIOA-----LNGQVLAEXTLVNOASTREPOVEA 7326
 Qy 1399 LENLKEDNNKLKEQAEESSEKQNSLEEFV---SGSQKLVDELVEYLKAQOLKAAEERLE 1454
 Db 7327 LQKAELEAMKALTEINKE-QIKASRYVNASDGLQANYNSALNGSOIATQPOPE 7385
 Qy 1455 IKDRYFELVQTANTLVEKLETPLOADHEEDSIDRSEEMEIKVLG-EKLERNOYLLE 1513
 Db 7386 LNKQVYNATQTIKT-----AENNLNGSKLAEKASDQNSQSTIEHLOGLTQ 7430
 Qy 1514 RLQEEKLELSNKLLEILO-----KEMETSVLLKDDLOQKLESLSSENIITLKENIDTT 1564
 Db 7431 SOKDKROHDLINQACTKQOVDDIVNNSKQLDNSM--NOLQOIV-----NNDNT 7475
 Qy 1565 LKHNSD-----TOAQLOKTOOEQLAKNLALIASDNCPTTOEKERSADCVHPL---EKKIL 1617
 Db 7476 VKQNSDFTINEDSSQDQAVNNAHIOAKKDLITA-----HPTINDKNOID 7517
 Qy 1618 LITTELHQNTNEOEKLEHKEKNELEQAOVELCEVEHLLKMSIESKSLSESIQHEKND-E 1676
 Db 7518 QALENIKALND---LHGSNKLSEDK-----KASQEQONLNSLNGOKDITIL 7562
 Qy 1677 QOLLAQKOQOVYTO--EKKELQOTHEHLTAEVDHLKENIELGLFKNEAQOKTTKEOCL 1734
 Db 7563 NHIESAPTRQOVGEKIASAKOLNMTMKALRDSIADNNHIILOSSKFYNEDSEQONAYNOA- 7621
 Qy 1735 LENKEKLEQSOHRLOCEELIEPLKSLKDKESALETLKESEQOYVYNLOMEMEVMLEMEELK 1794
 Db 7622 VNKAKNITNDQTPYMANDELOSVLNEVKQTKDNL-HGDQKLANDKTAQATLNALNLTN 7680
 Qy 1795 NSQRTVI---AERQLODDLRESEVMSLETQDDLKQAEAL-----QOQKQVOELTQ 1845
 Db 7681 QAQNGLETFQVONSRSREQKVQVLANQNLDMAKKLDDALGNDAIQTSNYINEDTQ 7740
 Qy 1846 -----ISVQEKISILENOMLYNATVKTLESE-RDDLN-----OSKOHLESE 1887
 Db 7741 QVNEDEYTRGKNITVAEQTN--PNNASPTNINTIADKITEARNNDLHGVOKLQAOQOOSINT 7798
 Qy 1888 IETLSLSEKEEPLEQAEKDKADARKTIDITETKISNIEBOLQOATNLKETLYERESL 1947
 Db 7799 INQMT-----GLNQAKQKQOLN--OEIOQTQTRSEV-HOVINKAQLDMSMTLQROS 7846
 Qy 1948 IQCKEQLALNTEHLERETKSR---DLALGKMEQREDEANRVIALTERKSSLE-EQINEN 2003
 Db 7847 ITDEHEVQTSNYINETVNGQTAVNNAVDRKQIINGSNPT-----ANPLEVERAASN 7900
 Qy 2004 VTTLKKEG-EGEKETFYLDQPSKOSSSQMEELRESLAKTKDLOLEBAKEVEISAT--NEIK 2060
 Db 7901 VKISDIALHGERE-LNDKNSKTRFAVNLDMNLQNA-----QKEALTHEICQATVISOVN 7953
 Qy 2061 NLTKAISLEBEILLONASILNEAVSERENLRHSKOQLVSELOSLTKLSRDA-----2114
 Db 7954 NIYNKAKALNDMKK-----LNDIYAQOVNVAQSNNTYNIEDSTPQNMVYDITLNHQAOSIIDQ 8009
 Qy 2115 FAQSKREKDEAVNKIASIAEBIKILTKEMDEFRDSKE-----SLOQSSHLEELCTY 2167
 Db 8010 VANPTMSHDELEMANINNIKALIMLDGE-HKIQQAKENAVANLLINSYNDLNPQORAIURL 8068
 Qy 2168 KTEILOMLKQKQEDINN--KLAEKYEVDELLOHLSLKEQLODQ-----QNELNREKLRN 2220
 Db 8069 VNEAQOTREKVAEBOLOSQAOLANDAMKHLNRNSIONOSVQESKRYINASDAKREOYNHAAVRE 8128

```

QY 2221 YE--LCERMDIMEKESVLRMQNEPQOEDVDARMDILSRNOEIOEMEKISAYSE 2278
Db 8129 VENINQHPFLDKEL--IKQITDGVNQNANL--NGVELLDADKNANOSIPTLHMLQA 8185
QY 2279 QHTLSSLSSELQKTEAH-----KHCMLNKESL-----SSTLSRSGSIQT 2321
Db 8186 QONALNEKINNAVTEEVAALIGAKLLDHAMENLEESIKKEQYKOSSNYNEDSDVOE 8245
QY 2322 EHVKLNTLOTLNFKKVVYRTAAVKE--DHSL--IKDYEKDLAAROKHDELRL--OLQ 2375
Db 8246 TYDNADVHTTELN--QVYNPLTSLIEDIEHAINENYNAOKKOLRGOKIYOTIDLADKELS 8303
QY 2376 CLEOGRKWSASAEELKFCF-----IEFLWELLFRKANITOS--V 2414
Db 8304 KLIDLTSSOSSSNQITACTARTREVAQAIEKASLNHAMALNKVYNKADVLSSNFI 8363
QY 2415 ODDSEYOVF--LNQVS-----TLOEELHKKGFPMQMLEEFDLHVDAKL 2459
Db 8364 NEDPEKKAAYOALNHVDSIIHROTNPMDPTVINSTIHELETAO--NHLGQ--OKL 8417
QY 2460 SEGMOQENRIASTIOLTKRLKAVOSKIOREITVYLNQPEAKLOEKKQEKELMRME 2519
Db 8418 AHAQODAAVYNGLIHL-----NVAQREYAINNT--NATTREKVAKNLNDNAQLD 8466
QY 2520 HHGFSASVMEENARILGILTVODESKKLOSRKML--ENELNLYKDDAMHGEKVAI 2576
Db 8467 KAMELQOVVAHKNLNDSKYLNEDSKYQOQYDRIADABQLNQTNPYL--EPYKVDI 8525
QY 2577 LODKLISNAEELNAMOVLTKRKQDNLOA--MKEIENLOKMAKGAAPY--KEEIDNL 2632
Db 8526 VKDNVLA--NEKILFGAEKLSYDKSNANDEIKHMYLNNAOKOSIKDMISHALTEYKOL 8584
QY 2633 -----KTKVVKIEMEKIKYSKAT-----DOEIVALKSCLEDEKESG-- 2667
Db 8585 LQQAARILDEAKSLEDYQVYITDTLTPNTEASEDKKERYQVSHQAATI--DKINGSN 8643
QY 2668 -----LRRIKELRRA--QADNODTYCVPKYOYKASPTPYTCGGSGSIVSTALVLOSE 2720
Db 8644 VSLDQVQALQOLQASBNLDGDRVEBAKYHANOT-----IDOLTHLNSLQOQ 8692
QY 2721 KAALERELSHYKKRYHNLSTRMSSEDEKRYKAKSDAHSHTGSHGSPK----- 2772
Db 8693 TA--KESVKNATKLEELIATVYNNNAQALNKYMGLEDFINHADVEUSDNTRKDDKIT 8749
QY 2773 --TETTHGPTPERSEMPSLHLCSPKSSSTKRVVSPNSEIYSOLVSPGKTGM-- 2828
Db 8750 AYDALEHGO-----DIQKTATONETKOAL--QOLIYAEFTSLNGFERLNNHAR 8795
QY 2829 ----KHILSPKVGHLHKKRALSPRSE-----MPTOHVISPCKTGLHNLTESTLFDLSS 2880
Db 8796 PRALEYIKSLIEKINNAOKSALEDKVTOSHDLLEHLEHTVNEG----- 8836
QY 2881 PCQOQKVOENLNSPKGLFVFKSKSMPCPSOF-----FDNSKLGDFSEL--NTASND 2932
Db 8837 -----TJLNDIMGELANAVNN--YAPTASINYINDNLKDNFTQAINNARALN 8886
QY 2933 KSOAENMWYEA 2943
Db 8887 KTGQGNLDFNA 8897

```

```

; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (409)...(436)
; OTHER INFORMATION: Xaa = any amino acid
US-09-724-517-2

Query Match
Best Local Similarity 5.7%; Score 841.5; DB 4; Length 1279;
Matches 342; Conservative 244; Mismatches 537; Indels 259; Gaps 47;

QY 7 VKVCVRVRLIORBOGOANLQWAKAGNNTISQVDG--TKSFNFDREPFMSHSTQIYOEIA 65
Db 7 VKVAVRIRPLCKKALNHQVCVAVIRNSQOVLIGRDRVFTFDVFPKNTODEVYNTCI 66
QY 66 VPIIRSLQYNGTIFAYGOTSGETTYTMMG-----TPNSLGIIPQAIQVEFKIIOEIP 119
Db 67 KPIVLSLIEGNAVFAVYQGTGSGTYTIGGHIAVYVEGOKIIPRAIOEIFOSISEHP 126
QY 120 NREPLRVSWMEIYENETVTKDLCDRRKPKLEIREDFNRNRYVYADLLEELVMPBEHYIQW 179
Db 127 SIDNVAVSYIEYKELRDLLELETSMKDLHREDKGTIVYGAKECHVESAGEVMSL 186
QY 180 IKKEKNRHYGETKMDHSSRSHTIFRMIYESRDRNDPTNSNODGA--VWVSHLNL 234
Db 187 LEMGNARHGTQOMNEHSSRSHTIFRMIYESRDRNDPTNSNODGA--VWVSHLNL 243
QY 235 VDLAGSRAQOTAGFVRLKEGCHINSPLTIGQYIKLSD--GQAGCFYIRRSKLTREL 293
Db 244 VDLAGSRAQOTAGFVRLKEGCHINSPLTIGQYIKLSD--GQAGCFYIRRSKLTREL 303
QY 294 QNSLGNAKAVIITCTIPVS--FDELTSTLOFPASTAKVANTPHVNEVLDEALLKRYRK 351
Db 304 KDSLGGSAKTVMITCVSPSSNDESLNSIKYANRANRINKTPTVNFSPSSDR--IDEMER 362
QY 352 EILDLKQLE-----NLESSSETKAQAMAKEHTTOLLAETIKQLHKEREDRIWH 399
Db 363 EILDLKQLE-----NLESSSETKAQAMAKEHTTOLLAETIKQLHKEREDRIWH 422
QY 400 LTNIVVASSQESQ-----QDQRYKRRRT-----WAPKIONSLH 435
Db 423 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 481
QY 436 ASGVSDFDMLSRPLGNFESKKAKEFSDMPSPPEIDSVCTEFSDFDALSMDNSGIDAENW 495
Db 482 VHSVSPRYSLDRIFAGFRTRSOML--LGHIEQKVLHCGQSDMSD--EESGQKSGT 537
QY 496 LASKVTHREKTSLHOSMIDFGQISDVQFHD-----SKENOLQYTPKDSGM 543
Db 538 RCRSRSMWQKPDSCVSLVSLDTQDETOKSDLENEIDLKIDLOESQELNLQKILNSERIL 597
QY 544 AECRKASFEKEITSLOQLOSKREEKELVOS-----PELITAELE--EOLSVK 590
Db 598 TEAK-----QKRELITINIKMEDLIKELKTGNDAKSVSKQSLVLTLEHDAQKVE 652
QY 591 ----ANKLEWYVNS--REHISNAEVOTDEKEVVRKEMSVLGDGYNASNDLQDSVYG 644
Db 653 LIETQKQOLELEKNDSDVAMKVKILOKERRKKMDAKLAVOYLQKKQDSKTLASISTON 712
QY 645 KRSSSHDECIEHRK--MLBOKIYDLLEFLENLKKSEENDKQSSQEDFMESITOL-- 697
Db 713 EKRANLEKAVHMKYQKQIOLQKRLNEENKROLDVAVIRKQOKIKE-----IOLKTG 766
QY 698 CEAIMAEKANALAEELALMB-----DNFQNILLENETKREIADLESLEKQENOTNEFE-- 750
Db 767 QEBGLFKAKEDDADACNLKRRKSGFSIDHLQKLDQOKWMLDEVEKVLNQROLELEAD 826

```

QY 751 ILEKTOKEHEAOLJHEIGSLK-KLVENAEYNTN-----LEEDL----- 789
DB 827 LKREIYVSKKALLQEKSHLENKKLRSSQALNTOSLKISTRNLNLEQELSEKNVQLOTS 886
QY 790 --ETKTLKEODIOLAEIKRRADNLOKKVRNFDLSVSGDSEKICEE--IFOLKOSLSD 845
DB 887 TAEKTKISEQYEV---LQKEKODLOKRRHVDDEKLNKGRVLSPEEHVLTQLEGITA 942
QY 846 AEAVRDQKCEKSPLESENLEKEKMEDTSNMYNOKERKASLEFK-----OLETEKSN 898
DB 943 LEAAIE-----YRNESIQNRQSLRAS-FHNLISRGANYLEKLACLSPEIRTIILFR 993
QY 899 YKKMEADLOKELQSAFNEINYLNGLLAGKVRPDLISRVLEKKVSEFSKOLE-KALEKN 957
DB 994 Y-----FNKYNVNR-----EAERKQOLYNEMKMKVLERDN 1024
QY 958 ALENVYCLSEKFLPNEVECLKNOISKASEEIMLKQEGESASIIISKOETIMOESQO 1017
DB 1025 MYRELESALDHLK-----LQC-----DRLLTLOQKEHE---QKMOILLHHFKQO 1065
QY 1018 ILQLTDEVTHT-OSKVQOTEEOYLEKKMHND-----LEFKYIRNK-----SEADL 1063
DB 1066 DEEGIMEFEFKYEDKIQOLEKDLVFKYKTSRDHKKKKELVGEAIIHQIAPSEYOAGDG 1125
QY 1064 IREMENLAGTMESEVKIADTKHELEETIRDE-----QLHEKKYFQAMQTIPIPTL 1118
DB 1126 VIKPBG--GGMLSEELKMASPESKMLSGEREMOSSASLQTPNPQKLMEDIPELPT 1183
QY 1119 SDSL-PPSKLYEGNSODPIINDYHNLALATERNNINVCLETENSLKEQYIDINTOLO 1177
DB 1184 HSLAPSPGHMLGNEN-----KTETDNOFTKSHSLSSQIQVGVNGRLH 1229
QY 1178 SIQASIEKSDQOKKODEBEVKKLLMELKGLHDLSQSLKLEQLENEVEYEKLOT 1237
DB 1230 GTPVKLCKRELKRLQIS-----ALELSLRSSL--GVGISMADSIYVSRKRPD 1276
QY 1238 LQ 1239
DB 1277 LK 1278

RESULT 14
US-09-641-807A-2
Sequence 2, Application US/09641807A
Patent No. 6440731
GENERAL INFORMATION:
APPLICANT: Berand, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6440731el motor proteins and methods for
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/641, 807A
CURRENT FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1279
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (409)...(446)
OTHER INFORMATION: Xaa - any amino acid
US-09-641-807A-2

Query Match 5.7%; Score 841.5; DB 4; Length 1279;
Best Local Similarity 24.7%; Pred. No. 9,1e-34;
Matches 342; Conservative 244; Mismatches 537; Indels 259; Gaps 47;

QY 7 VAVCVARVRLDREGDQANLQMKAGNNTISQVYG-TKSFNDRYFNSHESISQIYQELA 65
DB 7 VAVAVRIRPLCKEALHNHQCVRVPIPSQVILIGRDRVETDFVFGKNSQODEVYVNCI 66

QY 66 VPIISALQGVNGTIFAVGQTSCKTYTMG-----TPNSIGITPAIOEVFKTIQELP 119
DB 67 KPLVLSLEGVAYVFAAGTGGTITGGGIIASVVEGQGIIPRALQEIFOSISEP 126
QY 120 NREFLIRVSYMEIYEVETVDLLCDDRKKRPLIREDFNNRVVADLTELVAWPPHVIOW 179
DB 127 SIDFNKVSYLEYKEDLADLLELFTSMKMDLIREDKGNITVYAKKECHVASAEVWSL 186
QY 180 IKKEKRNRYGFTKMNDSHSSHTIFRMIVESRDRNDPTNSNCQGA-----VMVSHLNL 234
DB 187 LEMGNAARHTGTQOMHESSRSHAIPTTISICVHKHMEAAE--DGSWYSPHIVSKFHF 243
QY 235 VDLASERSAOTGAGCYRLAKEGCNINRSFLITGOYIKKLSD-GAAGGFYINDSLTRL 293
DB 244 VDLASERYTKGNTGERKESIQINSGLAGNYSALGDPRRKSSHPIYVDAKITRL 303
QY 294 QNSLGNNAKTVICTITPVS--FDETLSTLOPASTAKHVRNPHVNEVLDLALKRYRK 351
DB 304 KDSLGSATVMTICVSPSSNFDESINSLKYNANARNRNKPTVNFSPESDR-IDEMEF 362
QY 352 ETLDLKOLE-----NLESSETKAQAKEBHQLAEIKQHKREDRIWH 399
DB 363 EYKLREALQSOQAGVSOTTQINREGSPDTNRIHSLREQVALQGEEXXXXXXXXXXX 422
QY 400 LNTIVVASSQESQ-----ODQRYVKKRRVT-----WAPGKIONSLH 435
DB 423 XXXXXXXXXXXXXXXXXXXXXXXXXXXXGPPDKRIPYVPPFTHLGHYITYSR-QDSK 481
QY 436 ASGVSDFDMLSRPLPGNFSSKAKFSDMPSEPEIDSVCTEFSDFDALSMMOSNGIDAEN 495
DB 482 VHTSPMYSLDRILPAGFRRSQML-LGHIEQDKYLHQCFSDNSDD--EESGEGEKGT 537
QY 496 LASKVTHREKTLHOSMIDFGQISDSVQPHD-----SSKENOLYLPKDSGM 543
DB 538 RCRSRMWLOKPPSVSLVELSDPTQDETOKSDENEDLKIDQESQELNLOKLKNSERTL 597
QY 544 AECRASFEKETSLOOQLOSKSEKKELVOS-----FELKIALE--TOLSVK 590
DB 598 TEAK-----QKRELTINIKMKEDLIKELIKGNDAKSYSKOYSLAKVYKLEHDAQAYE 652
QY 591 ----AKNLEMYTNS--REHSINAEVQTDVEKEVYKRENSVLDGSGYNASNSDLODSVDG 644
DB 653 LLETQKQLELENKOLSDVAMKVLQKEFERKKMDAKLRYVLQKKQDSKRLASTLSTON 712
QY 645 KRLSSSHDECIRHK-----MLEQKIVDLEEFLENLKKSSENDKOKSSSEDPMESIQL--- 697
DB 713 EKRANLEQSVDMKRYOKIQLOKRLRENEKKEKQLDAYIKRQOKIKE-----IOLKTG 766
QY 698 CAIYAKKANALEELALNR-----DNFDIILENETLKEIADLRSIKENQETNEFE-- 750
DB 767 QEEGLKPRAEIDDACNLKRRKGSFGSIDHLOKLDQKKWLBEVKKVLANQOLELELAD 826
QY 751 ILEKTOKEHEAOLJHEIGSLK-KLVENAEYNTN-----LEEDL----- 789
DB 827 LKREIYVSKKALLQEKSHLENKKLRSSQALNTOSLKISTRNLNLEQELSEKNVQLOTS 886
QY 790 --ETKTLKEODIOLAEIKRRADNLOKKVRNFDLSVSGDSEKICEE--IFOLKOSLSD 845
DB 887 TAEKTKISEQYEV---LQKEKODLOKRRHVDDEKLNKGRVLSPEEHVLTQLEGITA 942
QY 846 AEAVRDQKCEKSPLESENLEKEKMEDTSNMYNOKERKASLEFK-----OLETEKSN 898
DB 943 LEAAIE-----YRNESIQNRQSLRAS-FHNLISRGANYLEKLACLSPEIRTIILFR 993
QY 899 YKKMEADLOKELQSAFNEINYLNGLLAGKVRPDLISRVLEKKVSEFSKOLE-KALEKN 957
DB 994 Y-----FNKYNVNR-----EAERKQOLYNEMKMKVLERDN 1024
QY 958 ALENVYCLSEKFLPNEVECLKNOISKASEEIMLKQEGESASIIISKOETIMOESQO 1017
DB 1025 MYRELESALDHLK-----LQC-----DRLLTLOQKEHE---QKMOILLHHFKQO 1065

Dd	304	KDLSGSAATVITLTCVSPSSSHFDELSLNTKYANRANIRNKPTVYNPSPSDR- IDEMEF	362
Qy	352	ELLDLKLQLE-----NLESSETQAQMAKEHTQOLIAETIKQLHKEREDRIWH	399
Dd	363	ELKLEALQSOAGVSGTTOINRGSPDTRIHISLEVOVALQEGEXXXXXXXXXXXXX	422
Qy	400	LNIVVASSQESQ-----OQORFKRRRYT-----WAPKIONSLH	435
Dd	423	XXXXXXXXXXXXXXXXXXXXXADGPARKPERPPTTPEPDTHLGHYIIPSR-ODSRK	481
Qy	436	ASGVSDPDMLSLPGNFSKKAFFSDMSPREIDSVCTEFSDFDALSMDSNGIDAENW	495
Dd	482	VHTSPMYSIDRIFAGFRTRSQML-IGHIEQDKVYLHCQFSDNSDD---EESGQKSG	537
Qy	496	LASVYTHREKTSLHOSMIDPQIDSVQFHD-----SKRNOQIYLPKSGDM	543
Dd	538	KRRSRWIOKDPDVSCLVELSDTODEYOKSPLENEIDIKIDCLQESDELNLQKLNSERIL	597
Qy	544	ABCRKASEFKEITSLQOOLQSGKEEKELVDS-----PELKIAELE---EOLSVK	590
Dd	598	TEAK-----QKKRELTINIKMKEDIKELITGDADKSVSKOYSLATYKLEHABAKYE	652
Qy	591	----AKNLEMTNS--REHSTNAEYQDVEKEVVRKEMSVLSDSGYNASNSDLQSDVG	644
Dd	653	LLETQKQLELENKDLSDVAKVVLQKEFRKKMDAALRYOVLQKKQODSKKLASLQION	712
Qy	645	KRLSSHDECEJENR---MEQITVDEEETENLKKNSXNDQKSESDPMFSITOL----	697
Dd	713	EKRAMELOSDVHHKYQKIQOLQKRLRENEKRKQDLAVITRQDKIKE-----IOLKTG	766
Qy	698	CEAIAEKANALIEELALMR-----DNFDNITLENETLKREIADLESKLKENQETNEFE--	750
Dd	767	QEEGLKPAEEDLDACNLKRRKSGFSGIDHOKLDEQKMWLDEVEKVLVNOQLELEAD	826
Qy	751	ILEKTOKEHHAQILHEISGLK-KLYNAEMKYQN-----LEBDL-----	789
Dd	827	LKRRRAIYSKKEBALQOESHLENKLLSSQALNDSLKISTRLNLEQELSEKNVOLQTS	886
Qy	790	--ETTKLKEQEIOLAELRRKRDNLQKRYNFDLSVMSGSEKLEE---IFOLKOSLSD	845
Dd	887	TALEKTKISEQEV---LQEKEDQILKRRHDVDEKLKNGRVLSPREHNLVQLEGICIA	942
Qy	846	AEAVYRQAKCSPSLRSENLLEKKMDTYSNWNQEKAKASLEK-----QLETKSN	898
Dd	943	LEAALTE-----YRNESIQNRQKSLRAS-FHNLSGEANYLEKACLSPVEITITLFR	993
Qy	899	YKKMEADLOKELQASAFNEIYNLNLLAGKVPBLLSRVELEKKVSEFSKOLE-KALEEN	957
Dd	994	Y-----FKYVYNLR-----EAERQOQIYNEMKMKYBLERDN	1022
Qy	958	ALNEBVTCLSTKYKLPNEVECLKANQJASKASEETMLLKQEGSHSASTISKOETIMOQSQ	1017
Dd	1025	MVRELESABHLK-----LQD-----DRRLTLOQKEHE---QKMOILLHHFRQ	1065
Qy	1018	ILQITDEVTHT-QSKVOOTEEOYLLEMKKMHD-----LFEKYIRNK-----SEARD	1065
Dd	1066	DGEIGMETFYETEKIQOLEKDLFYFKTTSRDKKKLLKELVGBALIRQLASPEYQDAGG	1125
Qy	1064	LREMENTKGTWSEYKVIADTKHELEETIRDK-----OLTHEKUYFQAMOTIPTIRPL	1118
Dd	1126	VLEPREG---GGMLSEELKWMASPRESMKLSGREREMSSASLRTQRPQKLMEDIDELRPI	1183
Qy	1119	SDSL-PPSKLVEGNSQDPIEINDYHMLALATPERNNIMVCLETENSLKEQYUIDLNTLOQ	1177
Dd	1184	HSLSPSGHMLNGEN-----KTEEDNOFTKSHSRLSOIOYVGVNGVRLH	1229
Qy	1178	SLAQOSLEKSDLOKPKQDLEGEVTKLLLEMLKGLHTDQSLSTIKLOLENEVETEKQJOT	1237
Dd	1230	GATPVKLCRKEMLKQIS-----ALELSLRSSL---GVGISMADSDIEVSKRPD	1276
Qy	1238	LQ 1239	
Dd	1277	LK 1278	

Sun May 25 14:36:17 2003

Search completed: May 23, 2003, 14:51:52
Job time : 77 secs

us-09-150-867-1.ral

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2003, 14:40:15 ; Search time 282 Seconds
(without alignments)
1038.851 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769
Sequence: 1 MSEGDAVKVCVVRPLQRE.....QAENMYEAKKETAPCKTS 2954

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1015	6.9	3899	9 US-10-171-311-4	Sequence 4, Appl1
2	1015	6.9	3917	9 US-10-171-311-8	Sequence 8, Appl1
3	1013	6.9	3907	9 US-10-171-311-2	Sequence 2, Appl1
4	1013	6.9	3925	9 US-10-171-311-6	Sequence 6, Appl1
5	993	6.7	3853	9 US-10-082-830-260	Sequence 260, App
6	821	5.6	2354	9 US-09-820-843A-113	Sequence 113, App
7	811	5.5	2310	9 US-09-991-496-120	Sequence 120, App
8	811	5.5	2310	9 US-09-820-843A-114	Sequence 114, App
9	811	5.5	2310	9 US-09-874-923-120	Sequence 120, App
10	794	5.4	2139	9 US-10-023-219-4	Sequence 4, Appl1
11	794	5.4	2139	9 US-09-727-384-6	Sequence 6, Appl1
12	761	5.2	1805	9 US-09-820-843A-73	Sequence 73, Appl1
13	720.5	4.9	2835	10 US-09-885-535-4	Sequence 4, Appl1
14	719.5	4.9	1979	9 US-09-927-597-4	Sequence 4, Appl1
15	718	4.9	6281	10 US-09-815-242-12996	Sequence 12996, A
16	717.5	4.9	1972	9 US-10-171-311-162	Sequence 162, App
17	708.5	4.8	1945	9 US-09-927-597-2	Sequence 2, Appl1
18	706.5	4.8	1938	9 US-10-171-311-164	Sequence 164, App
19	691.5	4.7	3158	10 US-09-815-242-12611	Sequence 12611, A

20	651	4.4	2025	10 US-09-815-242-5703	Sequence 5703, Ap
21	651	4.4	2437	10 US-09-815-242-5834	Sequence 5834, Ap
22	649	4.4	1103	10 US-09-847-874A-1	Sequence 1, Appl1
23	643.5	4.4	1786	9 US-09-742-096-3	Sequence 3, Appl1
24	624	4.2	2125	10 US-09-919-172-29	Sequence 29, Appl1
25	567	3.8	864	10 US-09-883-096-2	Sequence 2, Appl1
26	555	3.8	383	10 US-09-883-096-5	Sequence 5, Appl1
27	544	3.7	460	10 US-09-925-300-1228	Sequence 1228, Ap
28	527	3.6	338	10 US-09-883-096-4	Sequence 4, Appl1
29	526.5	3.6	1993	9 US-10-098-979-2	Sequence 2, Appl1
30	525.5	3.6	2285	10 US-09-932-183A-2	Sequence 2, Appl1
31	522	3.5	868	9 US-09-884-001-19	Sequence 19, Appl1
32	516	3.5	2055	9 US-10-017-216-4	Sequence 4, Appl1
33	514	3.5	5795	10 US-09-815-242-12610	Sequence 12610, A
34	500.5	3.4	2053	9 US-10-017-216-2	Sequence 2, Appl1
35	496.5	3.4	1591	9 US-10-197-668A-92	Sequence 92, Appl1
36	495.5	3.4	1958	12 US-10-028-946-4	Sequence 4, Appl1
37	495.5	3.4	2054	12 US-10-028-946-2	Sequence 2, Appl1
38	484.5	3.3	1203	9 US-10-097-340-43	Sequence 43, Appl1
39	481	3.3	1641	9 US-10-017-216-5	Sequence 5, Appl1
40	480.5	3.3	1597	9 US-10-017-216-6	Sequence 6, Appl1
41	477.5	3.2	2478	10 US-09-815-242-5816	Sequence 5816, Ap
42	477.5	3.2	2478	10 US-09-815-242-12967	Sequence 12967, A
43	474.5	3.2	890	9 US-10-171-311-196	Sequence 196, App
44	455.5	3.1	1368	9 US-10-159-151-2	Sequence 2, Appl1
45	447.5	3.0	3788	9 US-09-952-267-76	Sequence 76, Appl1

ALIGNMENTS

RESULT 1
US-10-171-311-4
; Sequence 4, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangt
; APPLICANT: Gialt, Karen
; APPLICANT: Gannavarepu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ. ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3899
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-4

Query Match 6.9%; Score 1015; DB 9; Length 3899;
Best Local Similarity 19.6%; Pred. No. 6.2e-31;
Matches 700; Conservative 630; Mismatches 1164; Indels 1072; Gaps 148;

QY 181 KKGGKNHYETKNDSSSHRTFRMIVESRRNDPTNSNCDGAYMVLNVDLAGS 240
DB 34 KKQKKR---KTSSSKHDVSAHDLNDIGOCNMYTINSSQRYESTVPIPESTIMRTLHSG 90
QY 241 ERASQTAEGVRLK-----EGCN--INRSLFLIGGVIRKLSDGQAGFTINRDSKLR 291

Db 91 EITSEOGFSEVLESEISTTADDCSSEYNGCSFYW-----RTGPTIN 132
 QY 292 ILQNSLGNKATVITICITTPVSPELTSTLOFASTAHVNTTPVNYVLDDEALLKRYK 351
 Db 133 LIREBERG-----VDSYS-----EGADDSPLHLMMSSELAGKOH-- 169
 QY 352 EILDJQKOLENLESSSETKA-----QAMAKEH---TOLLATIKOLHEREDRIMHT 401
 Db 170 EIELNLENEEMRYTGTGEGIQOJQEPFPAIKORDGIITQITANLQOARKREKDETFMEFL 229
 QY 402 NIYVASSQESQODRVKRRRTWAPKIONSLHASGVSD-----FDMLSR 447
 Db 230 EL-----TEOSQKIQIOFOOLQASSETLNRSTHSTAAIDLQAKQIILTHQOOLEBODH 282
 QY 448 LPGNFSK-----KAFSDMSPFEIDSCVCEFSDFDALMMSNGDAE 493
 Db 283 LBEYQKXKDEFTWQISFLOEKIKYEMEDCKVENSKEEIOEKETIIEIMTKIIEE 342
 QY 494 ---WNLASKVTHREKT---SLHQSMDPGQ-----ISDVOFHDSSKENOLQYLPKDS 540
 Db 343 KTLLELDKDLTTADKILGEIOEQIVQKNOEIKNMKLELTNSKQKERSSEIKQIM---- 398
 QY 541 GDMAEGRK-----ASPEKETS--LOOQLOSK--BEKKELVQSELTAELEBOL--SVKA 561
 Db 399 GYVBELOKRNHKSQFETDIVQREDETOKLOLRALDEMYGOOIVQMKOELIRIHOHA 458
 QY 592 KNEMYTNSR---EHSINAEVQDVEKEVVRKESYVLDGSGYNASNDLODSSVYDGRKS 648
 Db 459 QMEEMKTRHGKEMNLRYSNITVNDQIK-----LMNVAINELNIKLODTPNSQKREKL 513
 QY 649 SSNDECIEHKKMLQKIVDLEEFIEINLKNKSEMDKQSSBODW-----ESIOLCAPIMA 703
 Db 514 EEBGLILEECCALQOILEDVEELSFSREQIOARQITAEQESKLMEAHNSSTVEDLKA 573
 QY 704 EKANALE---ELAV-----MRDNFDNIIENETIKREIADLER----- 738
 Db 574 EIVASASRELELKHAETVNTYKIKLEMEKKNVLDOMASQO--AELEKRIOTLLS 632
 QY 739 ---SLKE-----NOETNEFEIL--EKETOKEHE 761
 Db 633 HEEELSKLEDELEIHRINIEKLKDNIGIHYKQOIDLQENMSQKLETWQFERDNLITRQ 692
 QY 762 AOLIHGSLKLVENAMYNOLLE-----DLETKTILKE-----QEIOL 803
 Db 693 NOILIEISKLKDQOS--LYNSKSEMTQIOLNELOKEITLROEKEKQTLDEQVEIOL 750
 QY 804 -----AELRRADNLQKVRNFDLSVS--MGDESEKLCBEIFOLKOSLSDAEAV----- 849
 Db 751 KTELLEKQMEKENDLOEKFQOLEAENSLIKDKETLEDMKIHHPVQSEERLIFIDSTK 810
 QY 850 ---TRDA--OKESFLASEN-----LEKKEKEDTSNMYNOKERKAASLEKOLETEKSN 898
 Db 811 SKSKDSVWKEKEITLLEENEDLKQOOCIOINEEIKORNTSEAKNFVAYOLOEYAC 870
 QY 899 YKMEADLO---KELQ-----SAFNEINYLN----- 921
 Db 871 LKVVDDLEDSKNQOLEKSKKALNEBLHQRINPTVKKSSVYEDDKTFVAYETLEM 930
 QY 922 GILAGKVPBDL-----SVALEKKVSEFSKOLEKALEKNALLENVOTLSEYK-- 970
 Db 931 GEVVEKDDTELKLEKLEVTREKRELSQRLSDSEOLKONGEISPLNEVYVSLKQEKQOV 990
 QY 971 -FLPREVELEKYN-----QISKASEIMLKQEG--EHSASITSK----- 1006
 Db 991 SIRCLELEITINHNAENVQSCDTQVSSILDDGVYMTSNGAGSVSKNKSSEGESKIMV 1050
 QY 1007 -----QEIIMOJQSEO---ILQTFDEYTHQS--KVQOTEOGYE-----MKKHNDLF 1050
 Db 1051 EOKVAFEMNTVGEESKQOLILIDHPSPYTKESLSLAPORSEMDKLOKELNVLKSEQNDL- 1109
 QY 1051 EKYINKSEAE-----DLIRB--MENLKGT-----MESVEVKIADTRKH 1086
 Db 1110 -----RLQMEQRICLSLVSTHVDQVREYNEKKAALCSLKEELIFQAEKIKELQKTH 1165
 QY 1087 ELE-ETIR-----DKBOLLH-----EKYFFQAMOTIPITPLSDSLPSPSKLY 1128
 Db 1166 QLEIQTMKQTOETGDEKRPHLILIGLKQKAVSECSYFLQICSV-----LGEYTPALMKC 1220
 QY 1129 EGNSDPIELINDY-----HNIALATEKNINMVCLET----- 1160
 Db 1221 EVNAEDKENSQDITSENEDPELODYREYQODPOENMHTLLNRYEYNLLVLIQRLSKI 1280
 QY 1161 -----ERNSLKEOV-----IDLNTLOLSQOASIKESDLOPK 1193
 Db 1281 WGOQTQGMKLEFGEENLPRKEETEFLSHSQMTNLBIDVNHHSKLSLQDLEKTYLEBY 1340
 QY 1194 QDLE-----EGEYKLLIEMELKGHL-----TDSQJST 1221
 Db 1341 QELESILSSLOOQKTEBQNYEAELHICLRQOAVSESTVPSLPVDSVYITESDAQRTM 1400
 QY 1222 -----EKLQENLEVTLEQILOEEMKNITP-----ERNE 1251
 Db 1401 YGSCYKKNIDGTIERSGFQKKEETINIVKLEKOQOELEEVAKVYSMSIAFAQOTE 1460
 QY 1252 L-----QTFEDLKAHDSILKQDLSENIEOSIETODELRAAQ 1288
 Db 1461 LSRISGKENTASSKQAHAVCOEQOHYFENMKLSODQIGQFFETVYD--VKREKEFKPLS 1518
 QY 1289 EELREO-KOLYDSFRQ-----OLDCSVGIS-----SPNHDAVANO---EYVSLGAVNS-- 1333
 Db 1519 KELGEHKGITILSNSDPHDIPESKDCVLTISEMFSKDTFTVROSIDELSVSSMDASR 1578
 QY 1334 ---LOSEMLGRDELDQSCALVSELELLR--AHVKSVEGENLEITKRLNGLEKEILKS 1389
 Db 1579 QLMNEQOLEDRORELVRQYQEHQOATELLROAHNRQHE-----RQ 1619
 QY 1390 ESESEVLKSLMELEKDNKTKQAEYSKENGQFLEEVFSGSQKLYDEIVKQKLA 1449
 Db 1620 REDQ-----EOLQEIIRKLNQOLAORSSIDENIVSE---REVLLEELKALQKSLAG 1670
 QY 1450 EERL--EIKDRDYFELVQNTANTNIVEGLKLETPLOADHEDSDIRSEMEIKVLGEKLER 1507
 Db 1671 REKLCCELRNSS---TQTONENQGEVE---EOTFEKELDRPREVPEILSNE-- 1720
 QY 1508 NOYLLELOEKELETSNKLEITQKEMETSVILKQD-----LOQKLESLSENIILKE 1559
 Db 1721 -RYALQOKAN---NRILKILLEVYKTTAAVEETIGRHVLGIIIDRSSKSSOASLWRS 1773
 QY 1560 NIDTTLK---HHSQOALQOK-----TOQOLOAKNL---ATAAS 1593
 Db 1774 EBAASVKGCVHEHRTVYVDESIPYSIGSDMRNDINMKSATERTETLSQRLVNSGFRGT 1833
 QY 1594 DNCPIQOE-----KETSAQVHPLEKILLLVTELOKQTEQEK 1633
 Db 1834 EIDPENEMILMINSRLQAAVEKILLEAISETSQOLEHAKVQTELMRESFQKQEAENSL 1893
 QY 1634 -----LHEKNEL--BOAOVELKQEVYEHLMKMSIESKSSLESLOHEKHD---TQOOL 1679
 Db 1894 KQOEELRERLHEESAPARQOLAVELS--KAEGVIDQYAEKTLFEQIOIEKTDIIDRLQOEL 1952
 QY 1680 LALQOQOVYQEKKELOQOTHEHLTAEYDHLKENIEIGLKNKNAQOKTKTEQOCLLENK 1739
 Db 1953 LCASNRLOELABEQOQIOEBELLISROKEMAK--AEAG-----PYEQOULOETE 1999
 QY 1740 ELEQSQHRLOEIEBELMSLKDKESALETILKESBOKYINLQOEMENYMLEMEELKNSQRT 1799
 Db 2000 KLMKEKLEVOQOAKVRDQLOQVKALE-----IDVEQYSR--FILEQOKNTELM 2049
 QY 1800 VIAREDQLODDLRSEVSEKSTFQODLKRAQEPALQOQKQVQELTSQISVL----- 1849
 Db 2050 DLROONALEKQOLKRMKRFQDEQAIDREHENDVQOQ--EIOKLEOQOLKVPBPQPSISEHQ 2107
 QY 1850 QEKISLENOMLVAVATVKEITLSERDLONSKQHLFSEITELISLKEKEFAL----- 1902
 Db 2108 TREVEOLANHLKEXTDKCSELLSKEQLOKQIOERNEEIKLEFRVARELEBOALLVEDRKH 2167

```

QY 1903 -----EOAKDKADARTIDITEK--ISNIEOILLQOATNKEPTLYERESL--IQCKE 1952
Db 2168 FGAVEANPELSLEVOLOAERAPADIKEREKITNEOELEOFREELNENEEVOOLMOLEI 2227
QY 1953 QALNTEHLEBETLKSIDLAKGME-----QOEDEANKVIALTEKMSLEE 1998
Db 2228 OKKSTIRLOELBOENKLFNDOMEKLGAIKESDAMSTQODHVLFGRFAOIIQCKEVID 2287
QY 1999 QINENVTTLKEGSEKETFYLOPSSKOQSSQOMELREBLSKTDLOLE-----E 2047
Db 2288 QLNKOQVYKLO-----QOLKITTDNKVIEKNEMLIDTLETOELCMLSDOECYKRN 2336
QY 2048 AEREISATNEIKNLAKTAKSISLEEIIQNASILNEAVSERENLHNSQOIVSE---LEO- 2103
Db 2337 REEIEOLNEVIEKLOELANIGOKTSMNASHSE---EADSLKHOLDVYIAEKLALEOQ 2393
QY 2104 -----LSLTLKSRDHAFAO-----SKREKDEAVNKIASLAE-----IKILTK 2141
Db 2394 VETANEMETPMKNVLFKFTNFKMNOLOJELPSLKRER--ESVFKIOSIPENSYNVAIDHLSK 2452
QY 2142 EMDEF-----RDSKESIQEOSSHUSE-----LCYTKTELQMLKQOKEDINKLAEKVK 2190
Db 2453 DKPELEVLTEDALKSLENQTYEKFSPENGKSGIINLETRILOESTVYSAKDELTOCYK 2512
QY 2191 EV-----DELLQ-----HLSLKEQ-----LDQIOMELANE----- 2216
Db 2513 QIKDMOGOFETEMLOKKTIVNLOKIVEEYAAALVSOIOLEAVOEAKFCQODNOTISSE 2572
QY 2217 -----KLRYVELCEKMDIMEKETISVLRLOQNEP-----QOEEDVAERMDILES 2260
Db 2573 PERTINIONLQNRDELGSDISALTIRLISLESQVEMHSLIIEKQVOLAER--NVLEK 2631
QY 2261 RMO-----EIOELMEKISAVYSEOHILLSSELQKETEPAHKHCMNLIKESLSTLSRSGS 2318
Db 2632 EKKLELEQKLEGENEKKOREKREKRSPODVEVLTETTELFF-----SNESGCFNELEA 2685
QY 2319 LQTEHVLTNLOLOLTKKFKVYVRYTAUVKEDH--SLIKDYE--KDLAEOK-----RH 2367
Db 2686 LRASVATKRELASYSKEKAELQOELVLYKETNMTSLOKDSQVNDHLAEKESILEKE 2745
QY 2368 DELALQ-----LOCT--EOHGRKMSDASEE---LKCEIEFLNELLFKRANTIQSVODFS 2419
Db 2746 DEFVEQSKKACMEPLPIKLSKISASQTDGTLKISSNQPOLVKNAGIQMLQSPCS 2805
QY 2420 --EYQVPLNOVGSTLOELEHKKFPMQLEFGDLHDAKLSGMOENRIST----- 2473
Db 2806 SEEVTEIISO-----FTEKIEKMOELHA-----ABILMESRHISSETTLK 2846
QY 2474 -----IOLLTRK---LKAVVQ-----SKI-----OREIT-----VY 2496
Db 2847 REHVAVOOLKEEGGTAKAYIOCLRSKSGSIPELASDAYSOTELICSSDSGSDMGOCIT 2906
QY 2497 LMO-----FEARKLOEKKEON-----KELMRMEHGPSA-----SYM 2528
Db 2907 LTHSGODIASEGGESESATDSFPKKIKGLRAVNHGEOVSLTESPYSDEGDSHQ 2966
QY 2529 E-----EENARLLIGITKTYODESKTIQSRKMLENELNLYKDDAMT-----KGEYAI 2576
Db 2967 QVSEPMLEERKAYINTISLKLDTKMQ-----LOREAEVYDSQSHESFSDMGDELILA 3021
QY 2577 LOKLKL-----SRNAEALNMOYKLFKQONDLAAMK----- 2609
Db 3022 LOQVTEERSVLLAAPTETLTALCTDAVGLNCLBOEIOGVEYQAMAECLOKADRRS 3081
QY 2610 ---EENIOMKVVAGAVPYKEEID-----NLTK-----VYKIEMEKIKYSAT 2650
Db 3082 LLSIQLALHOMGNRKITLKREGESEKPSQELLEYNIQKOSQMLEMGOVELSSMK--DRAT 3140
QY 2651 DQELAYIKSCLEDEEELRRLKEELRRARQADNDITVCVPKDYQASITFPYTCGGSGSIGVQ 2710
Db 3141 E-----LOEOLSEKKNVVAELKSELQTKLETTLKAQNHKL----- 3179

```

```

QY 2711 STAMVLQSEKAAERELSHYKKKYHHLSPRTSSSEDR-----KTKAKSDAHSHT 2762
Db 3180 -----ELEAFRLKVDKTDVEVHLNDLTSLAEQKSRLOALAEKREK-----L 3223
QY 2763 GSHRSGPKHTYETNRHG--PVTPERSEMPSLHSGSPKSSSTKRVASPNRSEIY SOLVMS 2821
Db 3224 GRSEERKEELEDLKFSLESQKORNLQNLLEQOKOLNNSQOKIESQRLYDAQJSEE 3283
QY 2822 PKGTGMKHILSPSKVGLHKRRALSPNRSEMPYQHVISPCKTGLHKLNTTESTLFDNLS SP 2881
Db 3284 QGRNLEQVLVLESEKRVLRREMSSTLDRERELHAQLOQSSDGTGOSRPLPSBLLKELOKQ 3343
QY 2882 C--KOQKVOENLN--SPKGLFDPVSK 2904
Db 3344 LEKHSRLVELNTEKTKKLDISQTR 3369

RESULT 2
US-10-171-311-8
; Sequence 8, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangl
; APPLICANT: Gialt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3917
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-8

Query Match 6.9%; Score 1015; DB 9; Length 3917;
Best Local Similarity 19.6%; Pred. No. 6.2e-31;
Matches 700; Conservative 630; Mismatches 1164; Indels 1072; Gaps 148;

QY 181 KKGCKNRHYGTCKKNNDSHRSHTIFRMIVESRDRNDPTNSGCGAVVSHNLNVLDSG 240
Db 34 KKQKKRR---KTSKSHDVSAHDLNDIOSQCNEMYINSSGRVSTYIPRESTIRTLHSG 90
QY 241 ERASQTAGEVRLK-----EGCN--INNSLFLGQVYKLSGQAGGFTINYDSKLR 291
Db 91 EITSHQGFVSYLESEISTADDCSSSEVNGCSFYV-----RTGKPTN 132
QY 292 ILQNSLGMNAVTYICITTPVSFDETSLDFASTAKHVRTPHVNEVLDDEALLKRYRK 351
Db 133 LTRREERG-----VDSYS-----EGQADSPTNLEMESESLAGKH-- 169
QY 352 IILDLKQOLENLESSSETKA-----QAMAKEH---TOLLAEIKOLHKREDRIWHLT 401
Db 170 EIEELNLEEMRYTYGEGLOLOEFEPAAIKQDGIITOTLNLQQAARRKEDTMEFL 229
QY 402 NIVVAESOQODRVKRRRYTAPGKIONSLSHAGVSD-----FDMLSR 447
Db 230 EL-----TEOSOKLQIOFOLOASSETLRNSTHSSSTRAADLLQAKOOILTHQOOLEEODH 282

```

QY 448 LFGNFSK-----KAFSDMPSPFIDSDVCTERSDFDALSMDSNGIDAE 493
 Db 283 LLEDYOKKEDFTMOISFLQEKIKYEMBOCKVENSKEELOEKTETIIEENTIIIEE 342
 QY 494 ---WNLASKVTHREKT--SLHQSMIDFGO-----ISDVQFHDSSKRENOLOYLTPKDS 540
 Db 343 KKTLELKOKRTADKLLELOEIOVQKNQOETIKMMLELTNSKOKROSESEELIKOLM----- 398
 QY 541 GDMACSR-----ASFEKEITS--LOQOLOS--EERKELVOSFEIKALEBOL--SYKA 591
 Db 399 GIVELOLRNHNKDSOFETDIYOQMEQOTQRKLEQLRALDEMYGGQIYOQMOELIRQMA 458
 QY 592 KMLEVNTSR--EHSINAEVOTDYEKEVVRKENSGLDSGYNASNDLODSYDVKRLS 648
 Db 459 QMEKRTKHKGEMENALRSYSNITVNEQIK-----LNNVALINELNILOPTNSOKETKL 513
 QY 649 SSHDCIHRKMLBQKIYDLEFIENLKKSSENDKOKSSEDFM-----ESIQCEALMA 703
 Db 514 EBLGITLEKCALOQLLEDLVEELSFSRQOIORAKOTIABOESKLNEAHSKISTVEDIKA 573
 QY 704 EKANALE---EIAL-----MRDNFIITILENTEKREIADLER----- 738
 Db 574 EIVASSESRKELELKHEAVYNNIKILEMEKKNNAVIDRAESQE--ALEERLNTOLLFS 632
 QY 739 ---SLKE-----NOETNEFEITL--EKETOKREHE 761
 Db 633 HEEELSKLEDELEIHRINIEKLDNLGIHYKQOIGLONEMSOIKETMOPEKDNLTQK 692
 QY 762 AOLIHEIGSLKLEVENAMYNONLE-----DETETKILKE-----OEIOL 803
 Db 693 NQILIEISKLDLOOS--LVNSKSEPMIQTOLNELOKEIIRQEKERKGLIEQOEL 750
 QY 804 ---ALKRKADNLQKVRNEDLVS--MGDSEKICEEIPOLKOSLDAV----- 849
 Db 751 KTELLEKOKMEKENDLOKEFAOLEAENSILKDEKKTLEMDLKHTPYQOEERLITFLDSIK 810
 QY 850 ---TRDA--OKRECSFLRSEN-----LEIKEMEDTSNMYNOKKAASLEPEKOLETFSN 898
 Db 811 SKSDOSWMEKIEILIEENEDLKQOOCIOLENEIEKORNTESFAKNEVYVOELQOEYAC 870
 QY 899 YKKMEADLO---KELO---SAFNEINVLN----- 921
 Db 871 LTKVKDLEDSKKNQOLEYKSKIKALNEELHOKINPTVYKMSVDEKTEPYAETLEM 930
 QY 922 GILAGKAPROL-----SRVLEKRVSEFSKOLEKALEKNALENEVTCISEYK-- 970
 Db 931 GEVVEKOTTEIMKELEYTKREKLELSQRLSDLEQOLKQGEISFLNEEYKSLKOEKEQV 990
 QY 971 -FLPNEVECLKN-----QISKASEEITMLKQEG--EHSASITISK----- 1006
 Db 991 SLRRELEIILNHNRAVNSOCTOVSLLDGYVYMTSRGAEGSVSKVNSFGESEKIYV 1050
 QY 1007 -----OEIIMQOSEQ--TLOLTBEVTHQS--KVQOTEEYOLE-----MKRMHDLF 1050
 Db 1051 EDKVSFENMNVGSEKQOELILDLHPVTKESSILRATOPSENDKLOKELNVLSKQNDL- 1109
 QY 1051 EKYIRNKSSEAE-----DLRE--MENIKGT-----MESEVYKIADTKH 1086
 Db 1110 ---RLOMEKORICLSVYSTHVQVREYVEMENKRDALCSLKEELLIPAOEKKIKELOKIH 1165
 QY 1087 ELE-ETIR---DKBQDLH---EKYFFQAMOTIFPITPLSDLSPLPSKIY 1128
 Db 1166 QLELOKMTKQETDEGKPHLLGLKQKANVSECSYFLOGLCSV-----LGEYTYPALKC 1220
 QY 1129 EGSQOPRIELNDY-----HNLIALATERNNIMVCLT----- 1160
 Db 1221 EVNAEDKENSXDYISENEDEPELODYREYODFOENNMHTLNTKYTEYNNKLIVLOTRLSKI 1280
 QY 1161 -----ERNSLKEOV-----IDNTQOQSIOQASIEKSDLOK 1193
 Db 1281 WGOOTQMKLEBEBNLPKEFEFLSIHSOMTNLEIDIVNHSKLSLQDLEKTKLEOV 1340

QY 1194 ODLE-----EGEVKILLEMELIKGLH-----TDSOLST 1221
 Db 1341 QELSELISLQOOLKETEONTEAIEHLOLRQAVSESTVPPSLPVDVYTTESDAQRTM 1400
 QY 1222 -----EKLOENLEVETKLOTOOEKKNITI-----ERNE 1251
 Db 1401 YPGSCVKNNDIGTIEFSGEGVKEETNIVKILEQYOOLEEYAVKYSMAJAFQOTE 1460
 QY 1252 L-----QTNFEDLKAHDSLKODLSENIBQSIETODELRAAQ 1288
 Db 1461 LSRISGKEWYASSKQAAVQOEOHYFENEMKLSODQIGFOTREYV--VKFKEEPFLPS 1518
 QY 1289 EELREQ--KOLYDSFRQ-----QILDCSVGIS-----SPNHAVANO---EVSIGEVNS-- 1333
 Db 1519 KELGEHEKTELLSDPHDIPESKDCVLTISEMFSKDKTFYIROSITHDEISVSMDSNR 1578
 QY 1334 ---LOSEMLGERDELOTSKALVSELELR--AHVKSVEGENLEITRYKLNGLEKELIGKS 1389
 Db 1579 QLMNLEBQLEDMQELVROYQEHQOATELLRQAHMRQME-----RQ 1619
 QY 1390 ESEVYLSKMLENKEDNNKLEQAEETYSKENOFSLEEVYSSQKIYDEIVLAKOLKAA 1449
 Db 1620 REDQ-----EOLOEETKRLNRQLOARSSIDNENIVSE--REBVLELELEALKOLSLAG 1670
 QY 1450 EERL--EIKRDYFELVOTANTMLVEKLETPLOADHEEDSIDRSEEMIRVIGEKLER 1507
 Db 1671 REKICELRNS--TQONGNENGVE--EQTFKEKEDRKEDVPELISME--- 1720
 QY 1508 NOYLLERLOEKLELSKLEIYLOKEMETSVLKD-----LOOKLESLEENITLKE 1559
 Db 1721 -RYALQAN--NKLKILLEVYKTAAYABETIGRHVGLIGLIDRSSOSASLITMS 1773
 QY 1560 NIDTTLK--HNSDOAQOLK-----TQOELQAKNL--AIAAS 1593
 Db 1774 EAAVASKSCHHEHTVYTDSTISYSGSDMPRNDIMMSKYVEEGELSORLVRSGFAGT 1833
 QY 1594 DNCPIQOE-----KETSADCVHPLEEKLITLLELHOKTNOEKL 1633
 Db 1834 EIDPENEEMLNITSRLQAAVEKILBAISETSQSLBHAVYQOELMRSEFRQKQATEEL 1893
 QY 1634 ---LHEKNEL--EQAOVELKCEVEHIMKSMLESLSLEIOHEKND--TEOOL 1679
 Db 1894 KCOEELRERLHESRARBOELAVELS--KAGGVIDGYADETLEPEROIOETDIIDRLDEOL 1952
 QY 1680 LALKQOAVYTOEKKELQOTHEHILTAEVDLKENILEGLNFKNQAOQTKKROCLLENK 1739
 Db 1953 LCASNRLQOLEAQQOIOERELLSHQKEAMK--AAG-----PVQOOLQOTE 1999
 QY 1740 ELBQSOHRLOCEIEELMKSILKDESALFTLKSEQKVNLNOMEMVYLEMELNSORT 1799
 Db 2000 KLMKEKLEVOCCAKERVDRLOKQVKALE-----IDVEQVSR--FELBQEKNTIELM 2049
 QY 1800 VIAERDOLDDLESEVEMSIETODDLRKAQEAALQOQKQVOELQISVL----- 1849
 Db 2050 DLKQNALKEKOLEKRRKRLDQALDREHERVFOQ--ETQVLEBQALKVPRPQISEHO 2107
 QY 1850 QEKISILENOMLYNAVATKELISERDLNOSQOHLEFSEITELSLIKEKEPFL----- 1902
 Db 2108 TREVELOANHLKEKTDKSELLSKREOLORDIOENNEETIEKLEFVRRELOALVLEDRKH 2167
 QY 1903 ---EQAKRDADARKITIDITEK--ISNIEBOLLOQANLKEPILERSL--IQCKE 1952
 Db 2168 FGAVEAKPELISLEVOLOERDAIDRKEETINLEBOLQOFREELKNKEEVOOLMOLEI 2227
 QY 1953 QALANTHELRETIKSKDLALGME-----QERDEANKYALATEKMSLEE 1998
 Db 2228 OKKESTYRLQOLEQENKLFKQDMKELGLAIKESDAMSQODQVHLEGRKQOITQGEVAYID 2287
 QY 1999 QINENVTILKESGEKEFYIQRPSKQOSSOMEELRESLTKDIOLE-----E 2047
 Db 2288 QINBOVTKLO-----QOLKITTDNKYIEEKNEILIRLELOIQLCLMSDQECVARN 2336
 QY 2048 AKKEISEATNEIKNLTAKISSLEEEITLQNASILNNAVSERENLRHSKQOLVSE---LEQ- 2103

```

Db 2337 REBELEQJNEVIEKLOELANIGQTSMAHSLSE---EADSLKQOLVIAEKALAEQ 2393
QY 2104 -----LSLTLKSDHAFQ-----SKREDAVNRKIASLAE-----IKILTK 2141
Db 2394 VETANEEMTFMKNVLEKTFNFKNLOLQELFSLKRER-ESVEKIGIPENSVAVALDHLJSK 2452
QY 2142 EMDER-----RDSKESLOESHSLEE-----LCTYTELOMLKQOEKEDINNKIAEKVK 2190
Db 2453 DKPELEVVLTDALCKSLSENGTYFNSFEENGKSIINLETRLQLESTVSANDLELTQCYK 2512
QY 2191 EV-----DELLQ-----HLSLKEQ-----LDQIOMELRRE----- 2216
Db 2513 QIKMOEGOFETEMLOKKIYNLQKIVEKYAAALVSOILEVAOYKAFQODNOTISSE 2572
QY 2217 -----KLRYELCEKMDIMEKETSIVRLMONEP-----QOEEDVAERMILLES 2260
Db 2573 PERTNIQNLQREDELSDISALTLRISSELSQVYEMHTSLILEKEQVEIAEK-NVLEK 2631
QY 2261 RNO---EIOELMEKISAVYSEQHTLSSLSSELOKTEHAKHCKMLNKSLSSTLSRSGS 2318
Db 2632 EKKLELOKLEEGNEKKQREKKRSPQDVEVLKTTTELFFH-----SNESGFENLEA 2685
QY 2319 LQTEHVKLNTQLOTLNKFVYVRYTAAYKEDH--SLIKDYE--KDLAEOK-----RH 2367
Db 2686 LRASVATKALASVKEKAEKLOEBLYKETNMTSLQKDLQOVDRHLAEAKESLIEKE 2745
QY 2368 DELRLQ---LOCL-EQGRKWSDSASEE---LKFCELEFINELLFKRANTIQSVODFS 2419
Db 2746 DETEVOESKAKMEPELRPIKLSKSIASQDGLTKISSNQTPIQLVKNAGIQINQSCBS 2805
QY 2420 --EVOVFLNQGSLQLEELHKKGMQWLEFGDLHVDAKLSBMOENRIAST----- 2473
Db 2806 SEEVTELTISO-----FTEKIEKMOELHA-----AETLDMSHRISETETLK 2846
QY 2474 -----IOLTKR---LKAAYO-----SKI-----OREIT-----VY 2496
Db 2847 REHVAVOVLKEEGCGLTKAVIQCRLSKRGSSIPELASHDAVOTRILCSDSGSDMGQCIY 2906
QY 2497 LNO---FEAKLOEKKEON-----KELMRMEHGPSA-----SVM 2528
Db 2907 LTHSGFDIASEGGESESATDPSPPKIKILRLAVHNEGQVSLTESPYSDEGHDHSIQ 2966
QY 2529 E-----EENARLIGLITKYODESKLOSRIKMLENELNLWKDAMH-----KGEKVAI 2576
Db 2967 QVSEPMLEEKAVYINTLSSLDLTKMO-----LQREAEVYDSSOSHSFSDMRGELLILA 3021
QY 2577 LQDKRL-----SRNAEFLNMQVYLTKRKQDNLQAMK----- 2609
Db 3022 LQVFLERBSVILAAFTLALGTDAVGLNCLQRIQOGVYQAAHMCLOKADRRS 3081
QY 2610 ---ELENLOKNAVAKGAVPYKEID-----NKTK---VAKIEMEKIYKAT 2650
Db 3082 LLSEIOALHQAONGRKITLKROESEKPSQLELEVINIOKOSOMLEMOVELLSMK-DRAT 3140
QY 2651 DQELAVYKSCLEDEKEGRLRKEELRRAQADNDTPVCPKYOXKASTPVCGGSGSIGVQ 2710
Db 3141 E-----LQBOUSSEKMYAEKSELAQTKLELTTLKQKHKL----- 3179
QY 2711 STAMVLQSEKALARELSHYKKRYHNLSTRMSSSEDR-----KTKRASKDAHSSHT 2762
Db 3180 -----ELEAFRLVYKDKTDVHLLNDTLASBQKKSRELQWALEKKA-----L 3223
QY 2763 GSHRSGPHKRETYRHH-PTYPERSEMPSLHSGPKKSESTKRYVSNBEITYSOLVMS 2821
Db 3224 GRSEERDEKELEDLKFSLESQKORNLQNLLEQOKOLNSQOKIESQRLMYDAQLSSE 3283
QY 2822 PGKGMKHHLIPSPKYGHLKKRRLSPNREMPQTHVISPQGTGLKHKNLTESTLFDONLSSP 2861
Db 3284 QGRNLEQVLLSESEKVRKRREMSSTLDRERELHAQLOSDGIGQSRPLPSDELKELQKQ 3343
QY 2882 C--KQAKVOENLN--SPKGLTFDVSK 2904

```

```

Db 3344 LEKHSRIVELNTEKRYKIDSLQTR 3369

RESULT 3
US-10-171-311-2
; Sequence 2, Application US/10171311
; Publication No. US20030087270A1
GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumel
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatz, Karen
; APPLICANT: Ganavathapu, Manjula
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-2

Query Match 6.9%; Score 1013; DB 9; Length 3907;
Best Local Similarity 19.1%; Pred. No. 7.4e-31;
Matches 691; Conservative 637; Mismatches 115; Indels 1182; Gaps 148;

QY 181 KKEGNRRYGETKMDHSSRHTTIFRMIVESDRNDPTNSCDAVAVSHLNLVDLGS 240
Db 34 KKQKKR---KTSSKHDVSAHHDNLNIDQSCNENYINSQVSESTVPESTIMTLHSG 90
QY 241 ERASGTGAEGYRLK-----EGCN--INRSLEFIIQVYIKLSDQACGFIYRPSKTLR 291
Db 91 EITSHEGFSVLESEISTADDCSSEVNGCSFV-----RTGKPTN 132
QY 292 ILONSLGNNAVTIICITTPVSFDETLSTLOFASTAKHVNRNPHVNEVLDEBALKRYK 351
Db 133 LLREBERG-----VDDSY-----EGQAQDSPTLHEMSESLAGKH-- 169
QY 352 ELIDKQOLENLESSESTKA-----QAMAKEH--TOLLAETKOLHKERDRIMILT 401
Db 170 EIEELNRELEKRYVYTGEGLOQLOEFPAJKRQGITTOGLANQARREDELMREFL 229
QY 402 NIYVASSQSOQDQVKKRKYTAAPKQIONSLSHAGVSD-----FDMLSR 447
Db 230 EL-----TEOSQKLIQFOQLQASFTLRNSTHSTADLILQAKOQILTHQOOLEQDH 282
QY 448 LPGNFSK-----KAKFSMPSPFEIDDSVCTEFSDDPDLASMDNSGIDAE 493
Db 283 LLEDYQKKEDFTMQISLQEKIKIYVEHQDKKYNSSKKEILOEKETIIEELNRTIIEE 342
QY 494 ---WNLSKVTTHREKT--SLHOSMIDFGQ-----HSDSVQPHDSSKENOLOYLPRDS 540
Db 343 KKTLELKDKLTTADLLLELOBOIVQKQEIKNMKELTETNSKQKROESSEETKOLM--- 398
QY 541 GDMAECR-----ASFKEITS-LQOQLOSK-EEEKRVLQSFELKIALDEQL--SVKA 591
Db 399 GTVELOQRNKHKDSQEFEDIQVQRMEOETQRLQRAELDEMYGQOIVQMOQELIRQMA 458
QY 592 KMLEKVTNSR---EHSINAEVQTDVEKEVVRKREMSVLQSDSGYNASNSDLQSSVDGRKLS 648

```

```

Db 459 QMEKTRHKGEMENALFYSNITVNEQIK-----LMNVAINEINIKIQTNSQKREK 513
Qy 649 SSHDECIEHRMLOKIDYDEEFIEINENKKSSENDKOKSSPDQF-----EIQCEALMA 703
Db 514 EELGILIEKALQODLVEELSFSPREQIARARQTAQESKINEAKHKLSTVEDIKA 573
Qy 704 EKANALE---ELAL-----MRDNFIILNENETLREIADLER-----738
Db 574 EIVASESRKLELKHAEVYNYKIEMLEKERNVAIDRAEAOE-AELERLATOLES 632
Qy 739 -----SLKE-----NOETNEFEL--EKETQREHE 761
Db 633 HEEELSKLEDELTEHRINIEKINIGIHYKQOIGDIGNEMSOIKETMOEKDNILTKQ 692
Qy 762 AOLHIEGSLKLVENAMYNQNLSE-----DLETKIKILKE-----QEIOL 803
Db 693 NQILLETISKLOLOS--LNSKSEEMTLOINELQKEITELROEKEKGTLEQVEQL 750
Qy 804 -----ALRRRADNLOKRYNFDLSVS-MGDSKLECEIFOLKQSLDAEAV-----849
Db 751 KTELLEKQMKENDLOEKFAQLBAENSILKDEKTLIEDMLKHTPYQOEERLITLDSIK 810
Qy 850 --TQDA--QKCSFLRSEN-----LEKEMDTSNWTNOKKASLEKOLETKSN 898
Db 811 SKSDSWWEKEIEITLIEENEDLKQOCIOLEIEKORNTSFBAKNEFVNYQELQEBEYAC 870
Qy 899 YKKMEADLO-----KELO-----SAFNEINYL-----921
Db 871 LKAYKDLEDSKKNQOELEYSKILKALNEHLQRIINPTVAKMSVDEDEKTPAETLEM 930
Qy 922 GLLAGKVPRLD-----SRVLEKVSSEKOLEKALEKNALNEVTCISEY--970
Db 931 GEVEKOTTEIMELKLEYTKREKTELQSRLDLSQKQHEISFLNEEYKSLQKQREOV 990
Qy 971 -FLUNEVECKN-----QISKASBEIMLKQEB-EHSAJISIK-----1006
Db 991 SLKRELEIITINNRANVOSCTQVSSLLDGVVNTSRAGEGVSXVNSFGEESKIMV 1050
Qy 1007 -----QEIIMQOSEQ--ILQITDEVTHQOS--KVQOTEQOYLE-----MKKMHDLF 1050
Db 1051 EDKVSFNMVTGVEESKQEOILDLHPVTKESSLRATQPBENDKLOKELNVLKSEQNDL- 1109
Qy 1051 EKYIRKNSAP-----DLRE-MENIKGT-----MESVEYKIADTKH 1086
Db 1110 -----RLOMEARICLSLVSTHVDQVREYEMENEKDALCSIKELIJAQEKELIKELQIKH 1165
Qy 1087 ELE-ETIR-----DKBQLH-----EKKYFQAMOTFFPTPLSDSLPSPKLV 1128
Db 1166 QLEIQTKTQETGDEGKPLHLILGKLOKAVSEBGSYFLQTLCSV-----LGEYTPALKC 1220
Qy 1129 EGNSSQDPIEINDY-----HNLATATERNINWCIET-----1160
Db 1221 EVNAEDKENGSDYISENEDELQDYRYEQDQENMHTLLNKTYEENKVLVYQTRLSKI 1280
Qy 1161 -----ERNSLKQOV-----IDLMTQLOSLQASQIESQDLQPK 1193
Db 1281 WGOQTQGMKLEFEGEENLPKRETEFLSHSOMTNLIEDVNHKSKLSLDLEKTKLEOV 1340
Qy 1194 QDLE-----EGEVILLLEMLLKGHL-----TDSQSLI 1221
Db 1341 QELIESLSSLOQOLKTEQOVYEAHICLOKRLQAVSESYPPSLPVDVYTTSQDQRTA 1400
Qy 1222 -----EKLQENLEVEKIQTLQOEBMKNTI-----ERNE 1251
Db 1401 YPGSCVKKNIDGTIEFSGEVKEETINIVKLEKQOEQLEEVANAVIYMSIAFQOQRE 1460
Qy 1252 L-----QTFNEDLKAEHDSIKQDLSENIEQSIETODELRAAO 1288
Db 1461 LSISSGKENTASSKQAHAVQOEOHYFNMKLSODOIGQPTREYD--VKFEKRPKPS 1518
Qy 1289 EELREO-KOLVDSFRO-----QLDQSVGIS--SPNHDAVANO--EKVSGEYVNS--1333

```

```

Db 1519 KELGHEGKEILLNSDPDHPIDESKDQVLTISEEMSKQKTFIYQSIHDEISVSMQASR 1578
Qy 1334 ---LQSMNGERDELQTSKALVSELELR-AHKSVEBENLEITKNGLEKELLGKS 1389
Db 1579 QLMNEBQLEDMQELVROYOEHQOATELLRQALMROME-----RQ 1619
Qy 1390 ESEVLSKMLNLEKEDNNKLEQAEYSKKNQFSLSEVSGSQKTYDEIVLKAOLKAA 1449
Db 1620 REDQ-----EQLOELIKRLNRQLOAQRSSIDNEVYSE--REVVILLEELKQSLQSLAG 1670
Qy 1450 EERL--BIKRDYFELVQANTNLVEGKLETPLOADHEEDSIDRSEMEIKVGEKLER 1507
Db 1671 RELKCELRNNS--TOTONGNENQGEV--EQFEKELDRKPEDVPEILSNE---1720
Qy 1508 NOYTLERLOEKELSKKLEILOKEMETSVLDD-----LOQKESLEBENITIK 1559
Db 1721 -RVALQKAN-----NKLKTLILEVNTAIVETIGRHVGLIDRSKQSQSSASLWRS 1773
Qy 1560 NIDTTLK--HNSDTQAOLOK-----TOOELOLAKNL--AIAS 1593
Db 1774 EAEASVSCVHEHTRTYDTSIPSYSGDMPRNDINMKSATVEGTELQSLVRSGFAGI 1833
Qy 1594 DNPCTIOE-----KETSAQVHPLEKILLLTELHQQTNEQEKI 1633
Db 1834 EIDPENELMNLNISRLQAAVEKLELAEISTSSQLEBAKTYQTEIMRESFRQOQATEST 1893
Qy 1634 -----LHEKNEL--EQAOVELKCEVHEIMKSMIESKSSLSLOHEKHD--TEOOL 1679
Db 1894 KQBELREHRESRAREQOAVELS--KABGVIDGYADEKTLFQRQOQEKTDIIDRLQOEL 1952
Qy 1680 LALKQOQVNTQOEKLEQOTHEHLTAEVDLKJENIEGLNFKKNAQOQTKTEQOCLINENK 1739
Db 1953 LCASNRLQOELVABQOQOIEERELLSRQKBAK--AEAG-----PVEQOLLQERE 1999
Qy 1740 ELQOSQRLQOCELEIMKSLKQKESALETLKESQVYINLNQMEVNLMEELKNSQRT 1799
Db 2000 KMKKELEYQOQAKERVDDLOKQVKALE-----IDVEEYVS-FLELEQEKNTIELM 2049
Qy 1800 VIAERQLODLDRESEVMSIETQDILKQAEALQOQKDKVOELTQOIS-----VLQ 1851
Db 2050 -----DIRQONALEKQLEKRRKRLDQALDREHERVYFQ 2085
Qy 1852 KISLENOMLYNAVYKTELSEKDLNOSKQHLFSEITLSLKEKE-----FALBO 1904
Db 2086 EIQKLEQOL--KVAPRFOPISE-----HQTRVEQOLANLKEKTKCSEFLSKBQ 2134
Qy 1905 AEKQKADAARKTIDIEKISNIEQLOQATNKEITLYRE--SLIOCKQOLAMT--EHL 1961
Db 2135 LQNDIOERNEIEKLEPRVARELQALLVSADTFQKVEDRKHGQVAKPELSLEVQLOAE 2194
Qy 1962 RETLKSQDLALQKQEOE--RDEAANKYALTEKKSLSLEQOINENVTTLKGEGERKEF 2017
Db 2195 ROAIDRKEKEITNLQEQLEQFRELENKKNBEVQQLMQOELQKKESTTLQOELQONKLF 2254
Qy 2018 YLQRPKQOSSQMEFLRSKTKDLOLEBAE-----KEISEATNFIKMLTAKISSL 2069
Db 2255 -----KQDMERLQALIKESQDAMSTQDOHVLFGRKFAQITQOEKEVYIDQINEQVTL 2304
Qy 2070 EEEIILQNASILNEVARENLHRSKQOVLSEQLSLTQKSDHAFQAKREKDEAVNKT 2129
Db 2305 QOOL--KITTDKVIIE-----KMLTRDLQTOEQCLMSQOECYKRRKEEIRQOLNE 2355
Qy 2130 A-----SLAEIKILTKMEDEFERDSKESLQOESSHISELCTYKT- 2169
Db 2356 IEKLOQELANIGQKTSMAHNSLSSEADSLKHQLDVYIAKLLAQOYVETANEMFMKNV 2415
Qy 2170 -----ELQMLKQOKEKEDINNKLAEKVKEVD-----LLOHSLKLEQLODQLOM 2211
Db 2416 LKETNFKANQLOLELFSLKRESEV--EKIQSIPENSVAVALDHLSKDKPDELEVVL 2470
Qy 2212 E-----LRNEK-LRNELECEKMDIMEKEISVRL-----MONEQ 2245
Db 2471 EDALASLENOQTYFKSFEENGKSTINLETRLOLESIVSAKULELTQCYKQIKQMOQOQ 2530

```

```

QY 2246 OEDDVABMDLIESRNOEIOELM-EKISAVYSEQHTLLSLSELOKET--EAKHCHL 2302
Db 2531 FE-----TEMLQKRYNLAKIVEEKVA-----ALVSIQLEAVOEYAKFCOD 2573
QY 2303 N-----IKE-----SLSTISRSFGSLQTEHVAKLNTLOTLINKKVVY 2341
Db 2574 NOTISEPERTNIONLREDELSDISALTLRISELESQVEMHTSL--ILKEQVEI 2631
QY 2342 RTAVKEHSHLKDYEKLAEQKRDHLRLOQLCEQHGRKWSASASELK----- 2393
Db 2632 AEKNVLEKELKLEQLKLEGNKKQREKE-----KKRSPQDVEVLTTELPHS 2681
QY 2394 -----FCEIEFL-----NELFRKANIIOSVODESEVOYL 2425
Db 2682 NEBSGFPELALRAESVATRAELASYEKAEKLOEBELLVETN--MTSLQDLDSQVRHL 2740
QY 2426 NOVGSSTIQ-----EELHKKG----- 2441
Db 2741 AEAKELSLIEKEDETEVOESKKAEMPEPLPIKLSKSIASQDTGLKISSNQTPQILVK 2800
QY 2442 -----FMQWLEBEGDLHVDAKKISEGMOENRRIAST----- 2473
Db 2801 NAGIOLNLOSECSSEVTEIISQFTEKIEKMOELHA-----AEILDMESRHISLETETLKR 2855
QY 2474 -----IOLLTKR-----LKAAYVO-----SKI-----OREIT----- 2497
Db 2856 EHYAVAVOLKEECGTLKAVIQLCKREKSSSIPELAHSDAYOTREICSSDSGDMGQGYTL 2915
QY 2498 NO-----FEAKLOEKKEQN-----KELMRMEHHGPSA-----SYWE 2529
Db 2916 THSOGFVLASGEGRSESATDFFPKIKGLRAVHNEGMOVLSTESPYSDGEDHSTIQ 2975
QY 2530 -----EENALLGLTKVVOESKLOSRKIMLEMLNLYVDDAMH-----KGGKVAL 2577
Db 2976 VSEPMLEKRAYINTISLKLDTIMO-----LOREAEVYDSSOSHESEFSPWRGELLAL 3030
QY 2578 QDKLL-----SHNAEAEILNAMOVLTKKODNLOAMK----- 2609
Db 3031 QOVFLERSVLLAARFELTALGTDAVGLNCLQRIQOEGVEQOAMECLOKADRSL 3090
QY 2610 -----EIEBNLOKNAVAKAVPYKEID-----NLTK-----VKTMEKIKYSKATD 2651
Db 3091 LSEIOALHACONNGRKITLKRQSESEKPOEILEYNIQOKOSOMLEMOVELISMK--DRATE 3149
QY 2652 QEIAVYKSCLDKEBGLRLKEELBRQAADNDTYVCVKDQKASTFFVTGCGSGIYQS 2711
Db 3150 -----LOQJLSEKKNVAELKSELAQTRLELETTLKAOHKLK----- 3187
QY 2712 TAMLVLOSEKALERELESHYKKKYHHLSTRSSSEDR-----KTKKASDAHSHTG 2763
Db 3188 -----ELFAFLRELVADKDEVHLLNDTLASEQKKSRELOMALEKKA-----LG 3232
QY 2764 SSHHSGPHKTYTRHG--PVTPERSEMPSLHLSGPKKSSSTKRVVSPNRSEIYQOLVSP 2822
Db 3233 RSEBRDKELDLKFLSLESQORNLQNLLEQOKOLNLESQOKLESQRLMYDALSEBO 3292
QY 2823 GKTGMHKKIILSPKVGGLKAKKALSPNSEMPTQAVISPGKTLKKNLTSTLFDNLSSPC 2882
Db 3293 GRNLEQVLESEKRYRIEMESTJDRERELHAQLOSSDGTQSGRPLPSDELKELQOL 3352
QY 2883 -----KOQKVOENLN--SPKGLKFDVKSK 2904
Db 3353 EKHSHRYELNETEKYKLDLSQTR 3377

```

RESULT 4
US-10-171-311-6

; Sequence 6, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan

```

; APPLICANT: Zhao, Xumel
; APPLICANT: Monahan, John
; APPLICANT: Kamathar, Shubhangi
; APPLICANT: Glatz, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersih, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3925
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-6

Query Match 6.9%; Score 1013; DB 9; Length 3925;
Best Local Similarity 19.1%; Pred. No. 7.4e-31;
Matches 691; Conservative 637; Mismatches 1115; Indels 1182; Gaps 148;

QY 181 KKEGNRHGTEKMDHSHRSHTIFRMIVESDRNDPTRENCDAVWVSHLNLVLAGS 240
Db 34 KKQKKR---KSSSKHVDYSAHNDLNDIOSQENEMVINSQVESTVPESTIMTILSHG 90
QY 241 EPASGTAGEVRLK-----EGCN--INRSLFIIQVYIKLSDQAGGFYWRDSKTLR 291
Db 91 EITSHOGRFVSELEISITPADCSSEVNGCSFVM-----RFGKPTN 132
QY 292 ILQNSLGNKAVIITCTPVSFDETLSTLOFASNAKHVRYNTPHVNEVLDEALLRKRYK 351
Db 133 LREBERG-----VDSYS-----EGQAQDSPTLHEMSESLAQKH-- 169
QY 352 ETLDLKQLENLESSSEKRA-----QAMAKEH---TOLLAETKQHKEREDRIWHLT 401
Db 170 EIEELNRELEEMRYVYTGEGLOQLOEPEAKIQORGIITQULANLOQARREDEMRREL 229
QY 402 NIVVASSQSDQVKKRRRYTAPGKIQNSLAHSVSD-----FDMLSR 447
Db 230 EL-----TEQSKRLQIQFOQLQASETLRNSTHSSSTAADLLQAKQILTHQQLREQDH 282
QY 448 LPGNESK-----KAKFSQMPSPFEIDDSCTERSDDDALSMDSNGIDAE 493
Db 283 ILLEDYQKKKEDFTMOISFLQEKIKYVEMEQDKVENSMBEIOEKETIIEELNTYIIEE 342
QY 494 ---NVLASVTHREKT---SLHOSMIDFGO-----ISDSVQFDSKENDLOVLPDPS 540
Db 343 KKTLELQDKLITADKLIGELQOIVQKNOEIKNMKLELTNSQKROSSEETKQML----- 398
QY 541 GDMACRK-----ASFKEITS--LQOOLQSK--EEEKKELVOSFELKIALEBOL--SYVA 591
Db 399 GTVEELQKRNRHDSQFEPDIYQRMQEQRKLEQJRAELDEYVGGQIYQMKELIRQMA 458
QY 592 KMLEVNTSR---EHSINAQVOTDVEKEVYKREMSVLDGSGYNASNSDLOOSSVUGKRLS 648
Db 459 QMEEMKTRHKGEMENALRSYSNITVNEQIK-----LNNVALNELINIKIQDINSQEKILK 513
QY 649 SSHDECIEHRKMLEQKIYDLEEFIEENLUNKSSKENDOKSSEODFM-----ESTOLCEAIMA 703
Db 514 BELGLIEBKALQORQDLVLEELSFSSRQIQRARQJTAQDSKUNEAHKSSTVEDIKLA 573
QY 704 EKANALE---EIAL-----MRDNFDNIIEENLTKREIADLER----- 738
Db 574 EIVSASESRKELELKHAEVYNTYIKLLEMLEKREKNAVADRMAESQF-AELERLRTQOLFS 632

```


QY 739 -----SLKE-----NOETNEPEIL--EKETOKENE 761
 Db 633 HEEBELSKIKEDLEIHRINIKELKDNIGIHYKQOIDGLOENEMSOIKLETMOFEKDNILITKO 692
 QY 762 AOIIEHIGSKIKAVENAMONLE-----DEFTKILKE-----OETOL 803
 Db 693 NOILIEISKIKDLOOS--LVNSKSEEMTLOINLOKEIETILROBEKEKGLLEBOVQLOL 750
 QY 804 -----AELKRADNLOKAVRNFDSVS--MGSEKLOEETIPOLKOSISDAEAV----- 849
 Db 751 KTELLEKOMKEKENDLOEFAOLEAENSILKDEKKTLEDMLKIHTPVSOEERLIFDSIK 810
 QY 850 --FRDA--OKECSFISEN-----LEKEMKEDISNMVNOKEKASLPEKOLEPEKSN 898
 Db 811 SKRSDWEKEIETILEENEDLKOOCIOLENEIEKORNTSPAEKNEVYQLOEETAC 870
 QY 899 YKREMDLO-----KELO-----SAFNEINYLN----- 921
 Db 871 LKAVKDLEDKSKMOELEYKSKIKALNEELHQRINPTVKMSVDEDEKTFVAETLEM 930
 QY 922 GLLAGVPRLL-----SRVELEKVSSEFSKOLEKALEKNALENYTCLSEYK-- 970
 Db 931 GEVEAKDTTELMEKLEVTYKREKLELSQRLSDLEBOLKOKHGEISFLNEEYKSLKOEKEY 990
 QY 971 -FLPNEVECLKN-----QISKASEEIMLKOEK--EHSASITSK----- 1006
 Db 991 SLKRELEIILINNRAENVOCDTOVSSLLDGVTMTSRAGESSVKNKSPGEEKIMV 1050
 QY 1007 -----OETIMOROSEO--TLOITDEVTHTOS--KVQOETEOYLE-----MKKMDLFF 1050
 Db 1051 EDKVSFENMTVGESEKQOELIDHLPSTVKESSLRATOPSENDKLOKELNVLKSECONDL 1109
 QY 1051 EKYIRNKSEAE-----DLBRE--MENIKGT-----MESVEKIKADTKH 1086
 Db 1110 -----RLOMEKORICLSIVSTHYDQVREYMEKOKALCSLKEELLIPAQEKIKELOKIH 1165
 QY 1087 ELE-ETIR-----DKEOHLH-----EKYFFQAMQOTFFPTPLSDISPSKLV 1128
 Db 1166 QLELOTKMQTOGTGDEGPHLLIGLKOKAVSEBSYFLOTLGSV-----LGEYTPALMKC 1220
 QY 1129 EGSOPFIEINDY-----HNLALATERNNIVCLET----- 1160
 Db 1221 EVNAEDKENSGLYISENEDELODYREYODFOENMHTLNTKYTEENKLLVLTQTLRSKI 1280
 QY 1161 -----ERNLSKEOV-----IDLNTLOSLQOASIEKSDLOPK 1193
 Db 1281 WGOOTGKMLEFEBENPKRETEFLSHSQMTNEDIDVNHKSLSSLOLLEKTKLEOV 1340
 QY 1194 QDLE-----BEGVALLMEMLKGLH-----TDSQLOST 1221
 Db 1341 QELESILSSLOOLKETQNYEAEIHCLQRLQAVSESTVPPSLPVDVITTESDAQRTM 1400
 QY 1222 -----EKLOENLEVTETKLOTLOEEMKNITI-----ERNE 1251
 Db 1401 YPSCVAKNIDGITTESGEGVKEETIVYKLEKQYOBOLEEBEYAKIVMSMASTAFQOTE 1460
 QY 1252 L-----QTNFEDLKAHDSLSKODLSENIETOSIETODELPAQ 1388
 Db 1461 LSRISGKENTATSKQAHAVCOQOEHYFNEMKLSQDIGFOTPEYVD--VKFKEEPFLS 1518
 QY 1289 EELREO--KOLVDSFRQ-----QILDCSVGIS--SPNHAVANO--EYVSLGEVNS-- 1333
 Db 1519 KELGEHKEKELILNSDPHDIPESDCVLTISEEMFSKDKFTIVROSITHDEISVSSMDASR 1578
 QY 1334 ---LOSEMLGERDELOTSKALVSELELR-AHVKSVEBENLEITTKLNGLEKEILGKS 1389
 Db 1579 QLMINEBOLEDMKQOELKYQOEHQOATELIRQAHMRQME-----RQ 1619
 QY 1390 ESEVLSKMLKENTKEDNNKLEQAEYSSKQENQFSTEEVSGSKIVDEIVLKAOLKAA 1449
 Db 1620 REDQ-----EOLQOEETIKRLNROLAORSSIDENLVSE--REVVILLEELKALQKSLAG 1670

QY 1450 EERL--EIKRODYELVOTANTNLVEGKLETPLOADHEEDSIDRSEEMEIKVIGEKLER 1507
 Db 1671 REKLCEELRNS--TOTONGENQGEV--EQFFKEKEDRKPEDVPPELISB-- 1720
 QY 1508 NOYLLERLOEKEKLESNKLEILOKEMETSVLLKD-----LOOKLESLENIILKE 1559
 Db 1721 -RYALOKAN-----NRLLKILLEVYKTTAAVETIGRHVIGLIDRSSQSQSSASLIWS 1773
 QY 1560 NIDTTLK--HSDIQAQOLK-----TQOELQALKNL--AIAAS 1593
 Db 1774 EASAVSKCVHEETRTVDESIPISYSGSDMPRNDINMMSKVTEEGTELQORLRSQFAGI 1833
 QY 1594 DMCPIQOE-----KETSADCVHPLEEKILLLETELHOXTNQOEXL 1633
 Db 1834 EIDPENBELMNLSSRLQAAVEKLELLEAISTSSQLHAYQOTLEHRESRQOAEETSL 1893
 QY 1634 -----LHEKNEL--BOAQVELKCEVEHLKMSMESSSLSLELQEKHD--TEOOL 1679
 Db 1894 KCOEELRERLHESRAREQALVELS--KABGVIDGYADEKTLFEROIOETKDIDRLQOEL 1952
 QY 1680 IALKOOMQVYTOEKLELOQTHEHLTAEVDHLKFNIEELGLNFKNEAQOKTKEQCLINENK 1739
 Db 1953 LCASNRLOELAEAOQOIOERELLSROKEMK--ABAG-----PVEQOLLOEHE 1999
 QY 1740 ELBOSORLOCEIEELMRSLKDESALETLKSEQKVININQEMEMVLMEEELKNSQRT 1799
 Db 2000 KLMKELEVQCAEKVARDLOKQVKALE-----IDEEQVSR-FIELEQEKTEML 2049
 QY 1800 VIAERQLODDLRESEVMSLETODDLRKAQOALQOKKQVQOELTSQIS-----VLOE 1851
 Db 2050 -----DLNOQOALEKOLEKRRKFLDQALDREHERVFOQ 2085
 QY 1852 KISLENOMLYNAVYKETLSEBDLNOGSKOHLFSEIETLSLSKEKE-----FALDO 1904
 Db 2086 EIOKLOQOL--KVPRFQPISE-----HGTREVOGLANHLKKTOKCSLELSKQ 2134
 QY 1905 AERKADARKATIDITEKSNIEBOLLQATNIKETLYERE--SLIOCKEQLAINT--EHL 1961
 Db 2135 LQRDIOERNEIEKLEFRVARELOALLVSDPTQKEDRKHGAVAKELSLLEVOLQOE 2194
 QY 1962 RETLKSQDALGKMOE--RDEANKVYALTEKSSSLEQOINENVTTLKEGEGKEFT 2017
 Db 2195 RDAIDRKEKEITNLEOEBQFRELEKNEVQOLMQLELOKKESTRLOEBOENKLF 2254
 QY 2018 YLOPFSKOSSQMEELRESLTKDLOLEBAE-----KEISEATNIEKNTLAKISSL 2069
 Db 2255 -----KDMERKIGLAIKESDAMSTQDQVHLEFGKFAQIOEKREVELOLQNEQVTXK 2304
 QY 2070 EEBILQNASILNEAVSREHLRHSKQOVLSELQSLTLKSRDHAFNOSKREDEAVNKI 2129
 Db 2305 QOOL--KITTDKVIIE--KNELLIDLETQIECLMSDQECYKRNEEIELOINEY 2355
 QY 2130 A-----SLAEIKITLREKENDERDSKESLOEBSHLSBELCYTKT- 2169
 Db 2356 IEKLQELANIGQKTSNMNHSLEBADSILKHQDYVIAEKLAEQOQVETANEEMTKNV 2415
 QY 2170 -----ELOMKOQKEDINNKLAEKVKEDE-----LLOHLSLKQDLOQIM 2211
 Db 2416 LKETNFKMNLOTOELFSIKRERSV-----EKIOSIPENSVNVAIDHLSKDKPELEVLT 2470
 QY 2212 E-----LRNEK--LRNYELCKMDIMEKETSVLTL-----MONEPO 2245
 Db 2471 EDALKSLENOITYKFSBEENGKSIINLETUOLLOESTVSAKDELQCYKQIDMOBOGO 2530
 QY 2246 QOEDDVAERMDILERSNOEIOELM--EKISAVYSEOHITLLSSLSSELOKET--BAHRHML 2302
 Db 2531 FE-----TEMLOKRIYVNOKIVKEKVAA-----ALVSQIOLEAVQYAFACOD 2573
 QY 2303 N-----IKE-----SLSTLSFSGSLQTEHYKLTQLOTLNKKVYV 2341
 Db 2574 NQTISSPERTNIONLOJREDELIGDISALTIRISELESQVQVEMHTSL--ILLEKVOEL 2631
 QY 2342 RTAAVVEDSHLINDYEDKDLAABQKRHDELBLOQCLEQEHQRKMSDSASELTK----- 2393


```

Db 2632 AKNVLEKELKLEQKLEBENKORKE-----KKRSPDVEVAKTTTELEHS 2681
QY 2394 -----FCEIFL-----NELFKKANIISVODDFSEVOVEL 2425
Db 2682 NEESGFNELEALRAESVATKALASYKAKAKILOEELIVAKETN-MTSLQDLSQVRDL 2740
QY 2426 NQVSTIQ-----ELEHKKG-----2441
Db 2741 AEAREKLSLEKEDETEVOESKKACMEPPLIKLSKSIASQTDGLKISSNQTPQILVK 2800
QY 2442 -----FMOWLEFGDLHVDAAKLSEGMQOENRIAST-----2473
Db 2801 NAGIOMLOESCESEVEYTEIISQTEKIEKMOELHA-----AELIDMESRIISTETILKR 2855
QY 2474 -----IOLLTKR--LKAIVO-----SKI-----ORBIT-----VYL 2497
Db 2856 EHYVAVOLLEKECCTLAIVQCLSRKESGSSIPELAHSDAYQTRICSSDSGDMGQGIYL 2915
QY 2498 NQ--FEAKLOEKKEON-----KELMRMEHHPRA-----SYME 2529
Db 2916 THSGGFIAESGGESESATDSFPKIKIGLLRAVHNEGMOVLSTESPYSDGDHSTIQ 2975
QY 2530 -----ENARLLGLIKTVODESKIKQSRIMLENEILNVKDMH-----KGEKVAL 2577
Db 2976 VSEFWLERKAYINTISLMDLITKMQ-----LQREAEVYDSQSHEFSFDMRGELLAL 3030
QY 2578 QDKLL-----SRNAEELNAMOVLTKKODMLQAMK-----2609
Db 3031 QOVFLERSVYLAARFETALGTTDVAGLNCLEORIOEGVEYOAMELIQADRSL 3090
QY 2610 --ETENLOKVAVAKGAVPYKKEID-----NLKTK-----VVIEMEKIKYKATD 2651
Db 3091 LSEIQAHAQONGKIKITLKRQESKPSOELEYNIOQKOSOMLEMOVELSSMK-DRATE 3149
QY 2652 QETAYLKSCDEEGEGRRLKEELRAQADNDITVCYPKDYOKASTEPVTCGGSGIYOS 2711
Db 3150 -----LOEQLSSEKMYVAELKSELQTKLETTLLKQHKMLK-----3187
QY 2712 TAMLVLOSEKAAERELSHYKKKYNHLSTRWSSSEDR-----KTKAKSAHSSHTG 2763
Db 3188 -----ELEAFRLKVDKTDDEVHLLNDTLASPOKSRLOMALKEKAK-----LG 3232
QY 2764 SSHRSGPKHKTETRYHG-PVPERSEMPSLHSGPKKSSSTKRYVNSRSIYSOLVMSR 2822
Db 3233 RSEBRDKEELEDKFSLESOKORMLQNLLEQOKOLLNESQOKITESORMLYDQOLSEBO 3292
QY 2823 KGTGMHKLHILSPSKVGLHKKRALSPNRSEMPQIVHISPGKTGLHKNLTESTLFDNLSSPC 2882
Db 3293 GRNLELQVLLESEKVRIRREMSSTLDRERELHAQLOSSDGTGOSRPLPSBDLLKELQOL 3352
QY 2883 --KOQKVOENLN-SPKGLFDVYKSK 2904
Db 3353 EKHRSRIVELLNETEKYKLDLQTR 3377

```

RESULT 5

```

US-10-082-830-260
; Sequence 260, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yonming
; APPLICANT: Recipon, Hervé
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Genes and Proteins Relating to Breast Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27

```

```

; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 260
; LENGTH: 2383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-830-260

Query Match      6.7%; Score 993; DB 9; Length 2383;
Best Local Similarity 21.2%; Pred. No. 2,66-30;
Matches 562; Conservative 472; Mismatches 929; Indels 690; Gaps 101;

QY 370 KAQMAKEEH--TQLLAEIK--QLHK-EREDRIHMLNIVASSQESQOQVRKRRRV 423
Db 227 KTOELKAEHRSOELIQLKSGODEKAELODRVETLSALLTQSOKEDEKIKALRE 286
QY 424 TWARGKIONSILHAGVSDPDMISRLPGNFSKAKAFSDPSPREIDDSVCTEFSFDDALS 483
Db 287 T-----VEILETNHTLMEHNASL- 305
QY 484 MMSNGIDAEMNLASKYTHREKTSLSHSMIDFGOI-----SDSVQPHOSSK 529
Db 306 -----SRNAOEKLSLOQYIKDITQVMYEEGDNIAQSGSHENSTEL-DS51 350
QY 530 ENOLOYLPKDSG-----DMAECKKAFSEKETSLOQQLQSKKEEK 569
Db 351 FSQFDYQDADKALLVRSVLTTRRQAVODLQQLAGCOEA-----VNLQOQHOMWEDEG 405
QY 570 KELVQSEFL-----KIAELEEOLSYKAKILEVNTNSRHSINAQVOTDVEKEVVR 619
Db 406 KALNRQLOKLTGERDITLAGQVVDLQGEVDSLSKEBELLOKARKE-----LR 451
QY 620 KEMSVLGDGYNAGNSD-----LQDSSVDGKRLSSHD--ECIEHRKMLEQKIVDLEEFTE 673
Db 452 QOLEVLEQEAURLRNVNVELOQDSAGQKEEQOEEHLVLRERERLOEMIMQLEA--- 508
QY 674 NLNKKSENDKQKSSQODEMESTOL-CEAIMAKANALEELALNDONITLLENTLKE 732
Db 509 ---KQSELSLITLRELESSEHLEGBELROEQIEVTRALRAQASIAELSSSENTLKE 565
QY 733 IADLERSLKENQETNEFEILEKETQKEHEAQLIHEIGSLKLIVENAEYNNONLEDELETK 792
Db 566 VADLRRAAVKLSALNEALALDQVGLNQQLQLEBENGVSGRMBAEGARALQVD--- 621
QY 793 TKLLKEQEIQLAEIRKRADNLOKRYRNFDSLVSQDSEKLCBEETFOLOKOSIDAEVATRD 852
Db 622 -----LAEAEKR-----REALMEKNTYHLEAQLQKAELEAGAE 652
QY 853 AQKESCFRSENLKKEKMEQDTSNMVNOKEKASLFEKQLETESNKKMKMADLOKELOS 912
Db 653 LQADLRDIOEKEKEIOKKLSER--HQQEAATTOLE-QLHQEAKROBEVLARAVOEKEA 708
QY 913 AFNEINYLNLGLAGVPRDLISRYELEKRYSEFSKOLEKALEE-----KNALEN 961
Db 709 LVREKALALEVRLOA-VERDQDQDLAEOGLSSAKELLESSLFEAQOQNSVIEYVKGQLEV 767
QY 962 EYVCLSEYK-FLPNEVECLAKNOISKASEEIMLLKQEGEHSATISKOETIMQOESQILO 1020
Db 768 QIOITVQAKVEYIOGEVRLCKLELDTL-----RSQAEOERDAARQ--LAQAQOEKTA 818
QY 1021 LTDEVITQSVQOOTEVOYLEMKKMHDDLFKEYIRNKSSEADLLREKMNLTGTMESYEVK 1080
Db 819 LEOQKAHKEKRVNQLREKMEKERSHQ-----QELAKALESLEREMLEEMR 865
QY 1081 IADTKHLELETIRDKEOLEHKKYFFQAMOTIFETPLSDSLSPSKLVEGNSODPIEIND 1140
Db 866 LKEQOTEMEAIOAQREBERPQAFSALCOMQ----- 895
QY 1141 YHNLIALATERNNIMVCLTEFRNSLKEQVIDNLTQDLSQAOSTEKSDLOKPKODELEGE 1200
Db 896 -----LETE-----KERVSLLETLLQTORELDAASOQLETRQDM--K 931
QY 1201 VKLLLEME---LKGHLTDSQLSIEKQLENLVETKIQTLQEMKNTTERNELOQNF 1257

```

```

Db 932 YOKKEBETGIIQTOQEORRELEKARQH---RDLALOESESSILDDKMDIQOVE 988
QY 1258 DLAEHOSLKDISENTEOSTIETODELRAOELEBOKLOVDSFRQOOLDCSVGISPNH 1317
Db 989 DLSQ--LVADDSQRLEVO-EVOEKRETOENRIKREL-----EREKASLISLMEK 1039
QY 1318 DAVANDERKSLGEVNSLOE---MLNGERDELQTSKALVSELELIRAHKSVGENMLET 1374
Db 1040 E-----QRLVIVQEADSTRQOELSRLRQDMQEOGECKELSAQMEHLLQVEKKEADPL- 1093
QY 1375 TKKLNGEKLILKSESESEVYLKMLLENKDKNNKLEQAEYKSEKNOFSLPEYFSSQK 1434
Db 1094 -----AOEQLLEBELSHTEQOL-----RASLMAOEAKAAQOLRLRSTES---- 1136
QY 1435 LVDEIEVLAQOLKAAEE-RLEIKRDYFELVOTANTNLVGLKETPLQADH-----EE 1486
Db 1137 ---QLELAAEQPGNAQOAOASLYSALOALSGVCSRPBLSGGGSASVWGLER 1193
QY 1487 DSIDRSEMEIYVGLERNOYLLERLOEKLEUS-NKL-ELLQKEMETSVLLKDDIQ 1544
Db 1194 DONGARS-----LEKRGPLLTALSAEAVASALHKLHODIMKTQOTRDVLRDQV- 1241
QY 1545 OKLESISENIIKENIDTLKHSDDQAOLOKTQOELQAKNIALASDNCPTQOKET 1604
Db 1242 OKLEERLTDEAKSOVHTELO--DIQROLSONQEKSR-----WEGKON 1284
QY 1605 SADC-VHPLKEKILLLEELHOQNTNEQEKLLHKNLELOAQVELKCEVHLMKSMIRKS 1663
Db 1285 SLSELMELHETMASLOSRLRRALQEMENAGREHLLQAKENITLQVHELOAAVVARA 1344
QY 1664 SLSLQHEKNDTEOQILALQOQNOVYQEKKE--LQOHEHLLTAVDHKNENIELGNF 1720
Db 1345 QASAGLTLEEDLTARSAKLKNEVESEEREAQLOEDELKYAQAQALQOENLAL- 1400
QY 1721 KNEAQOKTKEQCLNKENLEEOSQHLQCEIEELMKSLDKSALETLTKESBQVNLN 1780
Db 1401 -----LQOTLAREEVEETLRQOIQLEKQEMKAALELLSLDLK---RN 1444
QY 1781 QEKEMVWLEMEELKNSQRTVIAERDLODDLRESVEMSIETODDLKRAQALQOQNDKVO 1840
Db 1445 QEVDLQOEQIOLEKFC-RSVL-----EHLPAVQ-----EREQKLVQORQIR 1486
QY 1841 EL-----TSQISVLOEKISLLE-NOMLYNAVATKELISEBDDLNQSKOHLFSELETLS 1894
Db 1487 ELKRETOQRNVLEHOLELEKKQOMLE-----SQRGVODLKKQOLV-LECLALE 1536
QY 1895 LKKEKFALE-----QAEKKRADARKTIDITEKISNIEQ-----LLOQATNL 1937
Db 1537 LEENHHKMEQOQKLKELBEGORFORVALHLLDLERSQELQAOSSQIHDELSHTVL 1596
QY 1938 KETLYERESLQCK---EOLANTHELRKTLKSKDLAKGMOER-----DEANKVYA 1988
Db 1597 ARELQEDDOVKSQREQIEELQOKEHLODLERRQOEL-MQOKERIQVLEDRQOTKI 1655
QY 1989 LTKMSLSLEQINE---NVT-----LKESEGEKEFFLYORPSQOSSQOMEELRES 2037
Db 1656 LEEDDLQIKLSLREKRELTQOROLMOERAEBSG-----PSKAGQS-LEHKKLI 1705
QY 2038 LKTKDQLEAEKEISEATNEIKNLAKISLSEELIQNASILINENAVE---REMLRHS 2093
Db 1706 LRDKREKEVQOQHIELQELKQDQLOGLHRRKGEISLLSQREQELVYLQOOLQRA 1765
QY 2094 KQOLVSELQOLSTLKSRRHAFQASREKDEAVNKIASIAEELIKLTKEMDEFDRSKESL 2153
Db 1766 REQ--GELKEQSL-----QSOLDEAQRALAQORQOEALQOEQQAQOQGEER 1811
QY 2154 QBSSSHLSSEL---CTYKT---ELQMLKQOQEDINNKLAEKYKEVDELLQHSLSKQOL 2206
Db 1812 KEKADLOGLAQALQAHMTLREHGELODHKEQARLEELAEVEGRV-----QALEEVL 1864
QY 2207 DOJOMELRNEKLNRYELCEKMDIMEKEISVLRMON--EPOQE-----EDDVABERMD 2256

```

```

Db 1865 GDILRAESR-----EDEKALLIALQOQCAEOAEQEHVEVTRALQDSWLOQA 1908
QY 2257 ILSRNOEIOELMEKISAVYSEQHTLLSSLSSELOKRETEAHKQMLNKRSLSTLSRSF 2316
Db 1909 VIKEROEELALAE-----SOSKROEBAAR-----ABALDEALGAH 1949
QY 2317 GSIQTEHVKNLTQOLTLNFKVYRTAFAVEDHSLIKDYEKDLAEOKRHE-----LR 2371
Db 1950 AALQG-----KEOHLLEQ-----AELSRSLASTATIQ 1977
QY 2372 LOQOCLQHGKRRKSDASSELKFCETLEFLELFRKANIIQSVOODESEVQVNLNGYST 2431
Db 1978 ASLDACQASHRQ-----LEELRLIOEGEIDQDL-----RYQEDVQOLQALAQ----- 2021
QY 2432 LOEELERKKGFMQMLEFEGDLHADAKLSGMOQENRRIASTIQLTLKRLKAVQSKTOR 2491
Db 2022 RDEELRHQOERQDLER-----SLAQOVQENMTQOENQOGERE--EELRGLHQ--VR 2072
QY 2492 EITVYNQOFAKLQOEKQONKEIMRRMEHHGPSASVMEENARLLGLKTVODESKLOS 2551
Db 2073 ELQTLAQKQOELIETRETQORNNLEALPHSHKTSPEBESQSLKIDSLERPLQRELERLO 2132
QY 2552 RIKMLENE-----LNLVKDDAMHG--EKVALIQDKLISRNMAEELNANQVKL 2597
Db 2133 ALQOTEAREIEMREKADOLALSLAQTKASYSLQEVAMFLQASVLERDEEQ----- 2184
QY 2598 TKRQDNLOAMKEIENLOKAVAKAVPYKE-----EIDMLKTVVAKIEMEKIR 2645
Db 2185 -RLQDELELTRALKE-ERLHSPGATSTALSGRGQOVLQSVGSYEAEPSPDGMKOS 2242
QY 2646 YSKATQOELAYLKSLEDKEEGRLKLEELRRQADNDITVCYPKDYOKASTPVTYCGG 2705
Db 2243 W-----KORHEHLQOAAVARLEIDRSRLQRRN-----VOLNSTLE---ODG 2279
QY 2706 SGIVQSTAMLV--LOSEKALEREIS-HYKKXYHHSRTWSSSEDRKRTAKSDAHSHT 2762
Db 2280 RGQRNSDAKVALELOKEVYLLQNLQTLERKQODITYRSQOTRE-----L 2325
QY 2763 GSSHRSPPKFTETTYRRGPVTPPERSEMPSLHSGPKKSESTKRVASPNRSEIYSQLMSP 2822
Db 2326 AGLHSHLSHSLAVQAAP-----EATVL-----EATRRL-----DESLOQLTSP 2366
QY 2823 GKTCGMKHILSPS 2835
Db 2367 GPVLLHP---SPS 2376

RESULT 6
US-09-820-843A-113
; Sequence 113, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113
; LENGTH: 2354
; TYPE: PRT
; ORGANISM: L. major
; FEATURE:
; NAME/KEY: misc. feature
; OTHER INFORMATION: AC005802_5 16202.3
; NAME/KEY: misc. feature
; OTHER INFORMATION: g116899670
US-09-820-843A-113

Query Match 5 6%; Score 821; DB 9; Length 2354;
Best Local Similarity 19.4%; Pred. No. 9e-24;

```

Matches 514; Conservative 474; Mismatches 1013; Indels 646; Gaps 91;

```

Qy 333 TPNHVEVDDALLKRYKEIIDLK-----KOLENLESSSEFKQAAMKEHTOLLA--- 384
Db 123 SPQASVSTATLVHPEVEDAVSTKPSVSEADLHLSRIITFLQALNDEQINAAIATLS 182
Qy 385 EIKOLHKREDRIWHLTVIVVASSQ-----ESQODRVKRRVTAAPGKIQNSLAAS 437
Db 183 AAEOLRTAKEE-----NTALKSTAHLLQORLDTATQORAELEAVARALADRDRAAROL 236
Qy 438 GVSDFDLRLPGNFSKAKFSDMSPFEIDSVCTEESDDEALSMDSNGDAEMWLA 497
Db 237 AANNEELQORLDTATQORA-----ELEAVARALADRDRAAROLAAAN-----A 279
Qy 498 SKVHREKTSJHOSMIDEGQISDSVQFHDSKENGLOJLPDSCGMACRKAPEKETS 557
Db 280 EELQORLDTATQOR-----AELEAVARALADRDRAAROLAAANEE--- 320
Qy 558 LQOOLQOSKEEKEIVOSFELKIA-----ELEBOLSVKAKNLEMTNSREHSINAEVO 610
Db 321 LQORLDTATQORAELE-----EAVARALADRDRAAROLAAANEELOQRDLTA---TQOR 371
Qy 611 TDVEKEVVRKMSVVGSGVYASNSD--LQDSSVDGRKLSSSHCEIEHRKMLEOKIYDL 668
Db 372 AELEAQLARL-----AADRDRAAROLAAANEELOQRDLDTATQORAELEAVARL 420
Qy 669 EEFEENLKKSSENDKOKSSBODFMESIOLCEAIMAKANALEELALMDNDNTILEMET 728
Db 421 AANNEELQORLDTATQORAELEAVAR-----RLADRDRAAROLAAANEELOQRDLT 472
Qy 729 LKREIADLE-----RSLKENQETNEFEILEKETOKHEHAOLTHEIGSLKLVENAMYNON 784
Db 473 ATQORAELEAVARALADRDRAARQ-----QLAANEE-----E 504
Qy 785 LEEBLETKTKLLKQEOIOLAEKRADNLQKRVNFDLSVSKGDSSEKICEETPOLKOSLS 844
Db 505 LQORLDTATQORAELEAVARALAAANEELO-----QRLD 538
Qy 845 DAEAVTRDQACEGSLRESENLEKEMEDTSNMVYNOKKKAASLEFKEOLETEKSNYKMEVA 904
Db 539 TATQORAELEAVARALADRDRAAROLAAANEELOQRDLDTATQORAELEAVARALAAANEE 598
Qy 905 DLQOELOSAFNEINYNLGLAG-KVPRDLTSRVLELEKVFSEFSKOLEKALEKNALENEV 963
Db 599 ELQORLDTATQORAELEAVARALADRD--EARQOLAAANEELOQRDLDTATQORAELEAV 657
Qy 964 TCLS-----EKFLEPNEVECKAKNOISKASEIIMLLKQEGESASIIISKOETIIMQOSQOI 1018
Db 658 ARLADRDRAAROLAAANEELOQRDLDTATQORAELEAVARALADRDRAAROLAAANEEEL 717
Qy 1019 LQLDDEVYHTOSKVOOTEEOYLEKMKHMDLFEKYTRKKSABEDLREMEMLKGTMESEVE 1078
Db 718 QORLDTAT-----QORAELEAVARALADRDRAAROLAAANEELOQRDLDTATQORAELE 771
Qy 1079 VKIADTHELEETIRDKOEOLHKRYFFQAMQIIFPIPLSDSLPPSKLVEGNSODPIEI 1138
Db 772 AQLARLADRDRAARQOLAAANEELOQRDLTA-----TQORAELE 809
Qy 1139 NDYINMLALATERNNINWCVLETENRSLKEOYIDLNTLOLSIOAOSIEKSDLOKRODLEE 1198
Db 810 E-AQVAVARALADRDRAAROLAAANEELOQRDLDTATQORAELEAV-----VARLADRD 861
Qy 1199 GEVLLLEMLLEKGLHLDQSLSIEKLOENLEVTEKLOTELEEMKNITIEENELQTNED 1258
Db 862 ARQOLAAANEELOQRDLDTATQORAELEAVARALADRDRAAROLAAANEELOQRDLDTATQORAELEAVAR 921
Qy 1259 LKAHDSLKODLSENIOSIETODELRAAOEELREOKOLVVSFRQOLDSVSGISSPNHD 1318
Db 922 LAADRDRAAROLAAANEE-----LQORLD----- 945
Qy 1319 AVANQEKVSLGEVNSLOSSEMIRGERDELQTSCKALVSELELLRAHVSVSEBENLEITTKL 1378
Db 946 -TATQORAELE-----EAQOLARLADRDRAAROLAAANEELO-----QRLDTA--- 986

```

```

Qy 1379 NGLEKEIIGKSESEVLSMLENLEKDNKKLEQ-----AEYSSKENQFSLEEVSGSOK 1434
Db 987 -----TQORAELEAVARALADRDRAAROLAAANEE-----QO 1020
Qy 1435 LVDEIEVLKAOLKAAEEBLEIKDRDYFELVQYRANNLEVGKLETPQLADHEEDSIDRSSE 1494
Db 1021 RLDTATQORAELEAVARL--AADRDRAAROLAAANEELOQRDLDTATQORAELEAVAR-----QRA 1069
Qy 1495 EMEITVYLBKEERNOYLLERLOEKELELSNKLEITIQEKESVLLKDDLOCKLESLSSEN 1554
Db 1070 ELEAVARALADRD-----EARQOLAAANEELOQRDLDTATQORAELEAVARALADG 1121
Qy 1555 IILKENIDTT--LKHSDDTQAOLOKTOEOLOLANLAIASDNPITQOEKTSADCVHP 1611
Db 1122 DEARQOLAAANEELOQRDLT-----ATQORAELEAVARALADRDRAAROLAAANEE---E 1173
Qy 1612 LEEKILLTLEELHOKTNEOEKLLHEKNELEQAOVELKEVEHILMKSMIESKSSLES--- 1667
Db 1174 LQORLDTATQORAELEAVARALADRDRAAROLAAANEELOQRDLDTATQORAELEAVAR 1233
Qy 1668 LQHEKHDTBOULLA-----LKQOQVYTOBKKELOOTHEHLTAENDHLKENIEGLNFKNE 1723
Db 1234 LAADRDRAAROLAAANEELOQRDLDTATQORAELEAVARALADRDRAARQ--QLAANEEEL 1291
Qy 1724 AOKTKTEOCLLENKEKELEQSOHRLQCEIEELMKSL-----KDKESA 1765
Db 1292 QORLDTATQ-----QRAELEAVARALAAANEELOQRDLDTATQORAELEAVARALADRDRA 1347
Qy 1766 LETTESBOKVYNLQNEKEMVLEMEELKNSORTVYAEQDLODDRSEVSEKSTIOTDL 1825
Db 1348 ROOLAAANEE--LQORLDTATQORAELEAVARALADRDRAAROLAAANEE--ELQORL 1401
Qy 1826 RKAQALQOQDKVQELISQVLSQVLEKISLENGVLYAVATYKTELSEKRDLOKSGKHLF 1885
Db 1402 DTATQORAELEAVARALAAANEELOQRDLDTATQORAELEAVARALADRD--EAKOOLA 1458
Qy 1886 SEIETFLSLKEKEFALBOAEKDKADAKRTIDIREK-----ISNIEE--QLLOQATNLKE 1939
Db 1459 ANAELEOQRL--DTATQORAELEAVARALADRDRAAROLAAANEELOQRDLDTATQORA 1515
Qy 1940 TLYERESLI-----QCKBOLALNTEHLEETLSK-----DLALGKMEDEEANKVY 1987
Db 1516 ELEAVARALADRDRAAROLAAANEELOQRDLDTATQORAELEAVARALADRDRA-- 1570
Qy 1988 ALTEKMSLEBQINENYNTTLKEGSEKETFYIQ-----RPSQOQSSQOMELEARESLTKTD 2042
Db 1571 --ROOLAAANEELOQRDLDTATQORAELEAVARALADRDRAAROLAAANEELOQRDLT-- 1626
Qy 2043 LQLEBAKEKISATNEIKNLTAKISLSEEEILQNASIINEAVSERENRHSKQOLVSELE 2102
Db 1627 -----ATQORAELEAVARL-----AADRDRAAR--QOLAAANEE 1657
Qy 2103 QLSLTLSKRDHAFQASKREKDEAVANKIASLAEEIKILTKEMDEFKDSKESLOEGSHLSE 2162
Db 1658 ELQORL-----DTATQORAELEAVARALADRDRAAR-----QOLAAANEE 1696
Qy 2163 ELCTYKTELQHLKQOKEKEDINKKLEKVEVDELLHLSLSKEQDQIOMELRNEKLRYTE 2222
Db 1697 EL--QORLDTATQORAELEAVARALADRDRAAROLAAANEELO--QORLDTATQORAE 1750
Qy 2223 LCEKMDIMEKETISVLR-----LMQNPQOGEEDVAVARMDLTSRNOEIOELMKKI 2272
Db 1751 LEVEAVYLIRERERARGETAAGVEVOYLRTVEEVEEELCKEERKLESRAVQLR-----A 1806
Qy 2273 SAVYSEQHTLLSLSELQKTEEA-HKHCMNLKESLSSTLSRSFGSLQTEHVKLNLQLO 2331
Db 1807 SAAAKOQROEVAAKANEVEQERLDSMARCIAHGGAPOGADGRDQALROLANLREEVYLS 1866
Qy 2332 TLINKFKVYVYRT-AAVKEPDHSLIKDYEKDLAEOGRHDELQLOCLEQHRKKSDDSASE 2390
Db 1867 --EKOKAMERYITPGVRRORMLR-----AAEBORAD--LEARLVDEAGDLSRSPAA 1914

```



```

Db 1178 -----ELQORL---DTATQORAELEAQAARLAANAELQORLDTATQORAELEAQAARLA 1230
Qy 1670 HEKHDEQOOLIA-----LKQOQVAVTQOEKKEIQOTHEHLTAEDVHKENIELGLNFKNEAQ 1725
Db 1231 ADREBARQOOLANAEELOQRDLDTATQORAELEAQAARLAANAEELOQRDL-----TATQ 1284
Qy 1726 OKTKREOCLINENKELQOSQRLOCETEELMKSLKDKESALEFTLKESQKXININQOEKEM 1785
Db 1285 QRAELEBARVARLAADREBARQOOLANAEELOQR---DTATQORAELEAQAARL----- 1335
Qy 1786 VMLEMEELKNSQRTVIAERDQLODLRESYEMSIETODDKRAQEBALQOQKDYQOELTSQ 1845
Db 1336 -----NADGDARAQOOLANAEE-----ELQORLDTATQORAELEAQAARLAAN 1377
Qy 1846 ISVLOEKISLENOMLYNATVKTLEBRDDLNSKQHLSEITLSTLKEKEPALEQA 1905
Db 1378 AEELOQRDLDTATQORAELEAQAARLAADR---BARQOOLANAEELOQR---DTATQOR 1431
Qy 1906 EKDKADARAKTIDITEK-----ISNIEE--OLLQOATNKLKELYERESLI-----QCKEOL 1954
Db 1432 AELEAQAARLAADREBARQOOLANAEELOQRDLDTATQORAELEAQAARLAADREBARQOL 1431
Qy 1955 ALNTEHLRETLKSK-----DLALGKWEDEBDEANKVJALTEKMSLEQJINENYVTL 2007
Db 1492 AANAEELQORLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDLDTA 1544
Qy 2008 KEGGEKETFTLO-----RPSKOSSSOMERLRESLTKDLQLEAEKEISEATNEIKNL 2052
Db 1545 TQORAELEAQAARLAADREBARQOOLANAEELOQRDLT-----ATQORAE 1590
Qy 2063 TAKISLEEETILQNASILNEAVSERENLRHSKOOLYSELQOLSTLKSRRHAFQSKREK 2122
Db 1591 EAVARL-----AADREAR---QOLANAEELOQR----- 1619
Qy 2123 DEAVNKTASIAEETIKITKEMDEFDSKESIQEOSSHSELCTYKTELQMLKQOKEDIN 2182
Db 1620 DTATQORAELEAQAARLAADREBARQOOLANAEELOQRDLDTATQORAELE 1669
Qy 2183 NKLAKEKVEDDELQHLSSLEKEDDOJOMELNKLNYELCEKMDIMEKISTVLR----- 2238
Db 1670 AQLARLAADREBARQOOLANAEELOQRDLDTATQORAELEAQAARLAADREBARQETA 1726
Qy 2239 -----LMQNEPOQEDDVAERMIDLESNOEIQELMEKISAVSSEOHTLSSSELOK 2292
Db 1727 VAGQOQVLYREYVEEBCLEKRWKLESRAQLE-----ASAQKQOQOEVAANAAYOE 1782
Qy 2293 ETEA-HKCHMLNKESISLTSRSFSGSLQTEHVKLNTQLOTLNKFVYVYRT-AAVYEDH 2350
Db 1783 RLDSMARRCIAHEGDAPORADGRDALROLANLREEVYLS---EKOKAMERVIPGVBRO 1839
Qy 2351 SLIMDYEDKDLAEOGRHDELRLQOCLEQHGKRSKSDASSELKCEIEFLNELLFKKANI 2410
Db 1840 MRLE-----AAEQRAD---LEARLYDEAGDILSRPAST---NEVNLVYDLLOEH-- 1885
Qy 2411 IQSVODDFSEYOVLNOVGSTLOELEKHKGFQOMLEEGDILHDAKKLSGMOENRI 2470
Db 1886 -EAAQNNCITLEA---QVASTISDR-----DNGRQOESGDL 1917
Qy 2471 ASTIQLTLRLKAVOSQIQREITVYINQEFQAKQOEKKKONKELMRMHEHGPSASVME 2530
Db 1918 -----SEAOGRHLD-----NVQE-----RDMANH--RCALAE 1942
Qy 2531 ENARLLGLITVODESKKLSRIKMLNENL-----NLVKDAMHK- 2570
Db 1943 QNNAASELOAVKAKKLQASVYKASSLMTRLSASSGAGCVSARVRVGGSSAVPOAAHRD 2002
Qy 2571 GERVAIQLDKLRSNAEELNAMOVLTKKODNLQAAKETEINTQKAVAKGAVPYKEID 2630
Db 2003 AELIAEGERLRENGEMARLLAEGVELREBARPLEVIAE-----KLIGORSTDAEVA 2057
Qy 2631 NLKTKVYKIEKIKYKADQELAYL---KSCLEDEKEBGRILKEELR----- 2676
Db 2058 TEPTQV---RRNAASHRLDSREAOLEDERAARLREKEOQLLRVARELOQTSRALQVLYA 2113

```

```

Qy 2677 -----RAQ-----ADNDTYC---VPKDYQKASTFPVTCGGSGIVOSTAMLVLOSEKAA 2723
Db 2114 RALNRPOVTSLLTADGDISTYPPDTPOQOQGTPTP-----LREPVYS 2156
Qy 2724 LERELSHYKKKYHHLSTWSSSEDRKKTAKSDAHSSHTSGSHRSGPKHETYPHGPVTP 2783
Db 2157 LQSEYAH-----GRTAGAAV-----SSGLASPL-----PREP 2184
Qy 2784 ERSEMPSLHLSPKKSESSTRKRVSPN---RSEITYSOLVMSPGKTM 2827
Db 2185 PRARVYHVAEATGTEDTYQVRLTAATEARYDVLYEHLIESNGLOGV 2231

RESULT 8
US-09-820-843A-114
; Sequence 114, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PRO
; FILE REFERENCE: 063915
; CURRENT APPLICATION NUMBER: US/09/820, 843A
; CURRENT FILING DATE: 2001-03-30
; SOFTWARE: Patentin version 3.0
; NUMBER OF SEQ ID NOS: 118
; SEQ ID NO 114
; LENGTH: 2310
; TYPE: PRT
; ORGANISM: L. major
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: AC005893_12 L6202.3
; NAME/KEY: misc_feature
; OTHER INFORMATION: g1|6899664
US-09-820-843A-114

Query Match 5.58; Score 811; DB 9; Length 2310;
Best Local Similarity 19.48; Pred. No. 2,1e-23;
Matches 475; Conservative 450; Mismatches 926; Indels 596; Gaps 84;

Qy 558 LQOQLOSKKEKKEIYVOSFELKIA-----ELEEOQSVKRNKLEMYNREHSINAEVQ 610
Db 204 LQORLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDLTATQORAELEAQAAR 1708
Qy 611 TDVEKEVYRKEMSYLQSGYNASNSD---LQDSYVQGRKLSSHDECIEHRRMLEOKIVDL 668
Db 255 AELBARVARL-----AADREBARQOOLANAEELOQRDLDTATQORAELEAQAARL 303
Qy 669 EEFTEINLKKSEKENDKOKSSBEDFNESTQOLCEAIAEKAAMLEELAL---NRDFNDITL 724
Db 304 AANAEELQORLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDLTATQORAELEAQAAR 358
Qy 725 ENETLKREIADLESKLENOETNEFEILEKETQKEHEAOILHEIGSL-----KKLVE 776
Db 359 QRAELEAQAARLAANAEELOQRDLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDL 411
Qy 777 NAEYNTONLEEDLTKTKLKEDEIQLAEKLRADNTQKKVVRNFDLSVSMGDSKLCCEI 836
Db 412 NAE-----ELQORLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDLTATQORAELEAQAAR 452
Qy 837 FOLKQSLSDAEAVYTRDQKCEKSFLESENLEKEMEDTSNMVYNQKEKAASLFEKQLETEK 896
Db 453 BELQORLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDLTATQORAELEAQAAR 512
Qy 897 SNYKMEADLOKELQASAFNEINYLNGLAGVPRDLSTRYELKQVS-----EFSKOL 949
Db 513 ARLAADREBARQOOLANAEELOQRDLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDL 565
Qy 950 EKALEEKNALENETCYSEYFTLPNEVECLKNQISKAASELMILKQGEHSASTISQOEI 1009
Db 566 DTATQORAELEAQAARLAANAEELOQRDLDTATQORAELEAQAARLAADREBARQOOLANAEELOQRDL 619

```

QY 1010 IMOSEBOLITDEVTHTOSKVOOTEBOYLEMKMHDDLEFKYIRNKSEADILLREMN 1069
 Db 620 QLANAEELQRLDTAT-----QORAELEQRLADDEARQOLAAAEELQRLDT 673
 QY 1070 LKGTMESVEKADJTHELEETIRDEKOLLHEKKYFFOAMOTIPPIPLSDSLPSKIVE 1129
 Db 674 ATQORAELEQRLADDEARQOLAAAEELQRLDTA----- 713
 QY 1130 GNSODPIENDYHNLALATERNNIMVCLERNSLKEQVYIDLNTOLQSOAQSTF-KSD 1188
 Db 714 --TOORELE--AQRLAADDEARQOLAAAEELQRLDTATQORAELEAVARLAAD 769
 QY 1189 LQKPRDLEGEVYLLEMEELKHLTDQSITKLOLEMLEVTEKT-----QTLQEMKN 1244
 Db 770 GDEARQOLAAAEELQRLDTATQORAELEAVARLAADDEARQOLAAAEELQRLDT 829
 QY 1245 ITTERNELQTFEDLKAHDSLKODL--SENIEQSTIETODELRAA-----QOE 1290
 Db 830 ATQORAELEQRLADDEARQOLAAAEELQRLDTATQORAELEAVARLAADDE 889
 QY 1291 LREKOLVDSFROOLLDCSGVSSPNHDVANOEKVSLGEVNSIQSMLGENDERLOTSC 1350
 Db 890 ARQOLAAAEELQRLD-----TATQORAELE--EAQRLAADDEARQOL 933
 QY 1351 KALVSELEL-----LRAHKSVEGENLEITKRL-----NGLEKEILKSESEVIL 1395
 Db 934 AAANAEELQRLDTATQORAELEAVARLAADDEARQOLAAAEELQRLDTATQORAE 993
 QY 1396 KSMLENKKNKKLKEO-----AEYSKKNQFSLIEVFSQOKLVEIYVLAQOLKAAE 1451
 Db 994 EAQRLAADDEARQOLAAAEEL-----QORLDATQORAELEAVARLAAD 1037
 QY 1452 RLKIDRYELVOTANTNVEGKLETPLOADHEEDSIDRSEEMETKVLGKLEKRNQYL 1511
 Db 1038 RL-----AAAEELQRLDTATQ-----QRAELEAVARLAADDE 1072
 QY 1512 LERQOEKLELSKLETLQKEMETSVLLKDDLOKLESLSENITIKENIDTTLKHSHT 1571
 Db 1073 -----EAKQOLAAAEELQRLDTATQORAELEAVARLAADDEARQOLAAAEEL 1127
 QY 1572 QAOLOK-----TOOELOAKNLALIASDNCPIQOEKTSKDCVHPLEKI 1616
 Db 1128 EAQVAVARLAADDEARQOLADDEARQOLAAAEELQRLDTATQORAELEAVARLAAD 1177
 QY 1617 LILTEELHQTNEQKLEHKEKNELEQVLEKCEVETIKSM-----TESKSLESLD 1669
 Db 1178 -----EELQRL-----DVTATQORAELEAVARLAADDEARQOLAAAEEL 1230
 QY 1670 HEKHDPQOLLA-----LKQOMOVVTOEKKELOOTHEHLETAEVHLEKENTELGINKNEAO 1725
 Db 1231 ADDDEARQOLAAAEELQRLDTATQORAELEAVARLAADDEARQOLAAAEELQRLD 1284
 QY 1726 OKTTKEOCLINENKLEEOSQHRLOCEITELMKSLKDESALFETIKSEOKVINLQEMEM 1785
 Db 1285 QRAELEAVARLAADDEARQOLAAAEELQRL-----DVTATQORAELEAVARLA 1335
 QY 1786 VMELEMLKNSQRTVIAERDQLODDRESVENSIEFTODILKKAQOALQOKDKVOELTSQ 1845
 Db 1336 -----AADGDEARQOLAAAEEL-----ELQRLDTATQORAELEAVARLAAD 1377
 QY 1846 ISVLOEKISLENOMLYNATVKELESERDDNOSKOHFSELETLISLKEKEFALQEA 1905
 Db 1378 AELEQRLDTATQORAELEAVARLAADDEARQOLAAAEELQRL-----DVTATQOR 1431
 QY 1906 EKDKADARTIDITEK-----ISNIEE--OLLOQATNLKETIYERESLI-----QCKEOL 1954
 Db 1432 AELEAVARLAADDEARQOLAAAEELQRLDTATQORAELEAVARLAADDEARQOL 1491
 QY 1955 ALNTEHLRETLKSK-----DLALCKMEQDERDEANKVIALTEKSSLEQONENVTTL 2007
 Db 1492 AAANAEELQRLDTATQORAELEAVARLAADDEARQOLAAAEELQRLDTA 1544

QY 2008 KEGEKEFTFYLO-----RPSKOSSSQWMELERESLTKTDLOLEAKKEISEATNEIKNL 2062
 Db 1545 TOORAELEAVARLAADDEARQOLAAAEELQRLDT-----ATQORAELE 1590
 QY 2063 TARISSLEEEILONASTILNVAERENLRHSKOOLVSELEQSLTLKSHDHAFOQSKREK 2122
 Db 1591 EAVARLA-----AADRDEAR--QOLAAAEELQRL----- 1619
 QY 2123 DEAVNKIASLAEIKILITKMDERFSKESLQSSHLSLELCTYTELOMKQKQEDIN 2182
 Db 1620 DVTATQORAELEAVARLAADDEAR-----QOLAAAEEL--QORLDATQORAELE 1669
 QY 2183 NKLAKEKVEDELLLOHLSLKEQLODQLOMELREKIRNLECEKMDIMEKEISVLR----- 2238
 Db 1670 AQRLAADDEARQOLAAAEEL-----QORLDATQORAELEAVARLAADDEARQOL 1726
 QY 2239 -----LMONEPOEEDDVAERMDLIESRNOETLOELMEKISAVYSQHTLSSLSLELOK 2292
 Db 1727 VAGEQVQVLYREVEEBCLEKEERWCLESRYAQLRE-----ASANAQOQROEVAKANEVOE 1782
 QY 2293 ETERA-HKHCHLNTKESLSSTLSFSGSLQTEHYKLTQLOTLNKKVYRT-AAVKEH 2350
 Db 1783 RUDSMARCTIAHGDAPOQADGRDALROLANLREEVKS--EKQKAMERYIPGVREQO 1839
 QY 2351 SLIKDYEKDLAOKRHHDELRLQLOCLEQHGKWSDSASEELKFCIEFTNELLEFKKANI 2410
 Db 1840 MRLE-----AAEQRAD--LEARLYVDEAGDLRSRPAAT--NEVNLRYDLALQEH-- 1885
 QY 2411 IOSVODFSEVOVFLNQVOSTLOEELHKKGFQWQLEFGDLHVDKLLSEGMQOENRI 2470
 Db 1886 -EAAQNRCTTLEA--QVASLTSDR-----DNROQOESADL 1917
 QY 2471 ASTIQLTKRLKAVQSKLORETIVYLNQFEAKLOEKKQENELMRMHGHPASVME 2530
 Db 1918 -----SEARHHD-----INQD-----RDMANH--RCALAE 1942
 QY 2531 ENARLIGITVODESKLKQSLKMLNEL-----NLVKDDAMHK- 2570
 Db 1943 ONAAMASELOAVAKLRQASVYKASSIMTRLSSAGGAVSARVYGVSSAVPOAAPHRD 2002
 QY 2571 GEKVALLQKLSRMAEALNQVYLTQKQONLQAMKEIENLQKMAKGAVPKKEID 2630
 Db 2003 AELIAVGERLREKRGAMLLABGVELREARPLEVYLA-----KLIDRRTSAAEVA 2057
 QY 2631 NLKTVVKKIEMEKIKYSKATQDEIAYL--KSCLEDEKGLRLKEELR----- 2676
 Db 2058 TEPTQV-----RRMAHSRLDSREAOULDERAARLREKEQQLLRVARELOTKSRALQVLYA 2113
 QY 2677 -----RAQ-----ADNDTVC--VPKYQKASTFPYTCGGSGSGLVQSTAMLYQSEKRA 2723
 Db 2114 RALNRPQVTSLLTADGDTSTYDPDPQOQOQGTKRP-----LREPYYS 2156
 QY 2724 LERELSHYKKKYHHLSTRUSSSEDRKKTAKSDAHSHTGSHRSGPHKTTETRYRGVTP 2783
 Db 2157 LSEVAHY-----GRTGAAY-----SSGLASPL-----PREP 2184
 QY 2784 ESEMPSLHSGPKKSESSTKRVVSPN--RSEIYSOLVMPSPGKTGM 2827
 Db 2185 PRARVHRAVEATGTEEDTQVRLTATEAVRDLVYEHILLESNGLOV 2231

RESULT 9
 US-09-874-923-120
 ; Sequence 120, Application US/09874923
 ; Patent No. US20020081320A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Campos-Neto, Antonio
 ; APPLICANT: Webb, John R.
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Colter, Rhea

QY	2351	SLIDYEKDIAEQRKRDDELRLQLOOEQHGKRWSDSSAEELKPCOEIEMLELFFKANI	2410
Db	1840	MRLD-----AAEQRORD--LEARLVDEAGDLRSPPAST--NEVNLVRLALQEH--	1885
QY	2411	IOSVODDFSEYOVLNOVGSTLOOELEHKKGFQMWLEEFGLDVLDAKILSEGOENRI	2470
Db	1886	EAQNRCSTLEA--QVASTLSDR-----DNGROGESADL	1917
QY	2471	ASTIQLTKRLKAVQSOIRELTYYVLNOGPEAKLOEKKQKNELMRRKEHHGPSASVNE	2550
Db	1918	-----SEAGRHLID-----NVQD-----RDMANH--RCALTEE	1942
QY	2531	ENARLLGILTVODESKRIKOSRIKMLENEL-----NLVKDAMK--	2570
Db	1943	QNAAMASELQAVKAKILQASVKASLSMTRLSASSGAGSVARVAVGSSAVPQAAPHRD	2002
QY	2571	GEKAVILLQDKLISNNAEALNMAQVYLTKQDNQAAAMEINLQWAKAAGVAPKEID	2630
Db	2003	AELTAEGERIRREGEMRLLAGVELREARPLEVLALE-----KLGDRTSDAEVA	2057
QY	2631	NLKTGVVKEIEMEKIKYSKATDOEIAVL--KSCLEKEBGLRLKEELR--	2676
Db	2058	TEPIQV-----RRNAHSHRLHDSREAOULDERAARLKRKQGLLRVARRELQTSRMLQVYA	2113
QY	2677	-----RAQ-----ADNDTVYC--VKRQYKASTFPEYTCGGSGSIVOSTAMIVLOSEKA	2723
Db	2114	RALNRPOVTSILLTLTADDDTSYDPTQOQOQOGRTP-----LNEPYS	2156
QY	2724	LERELISHKKKYVHLHSLRTSMSSSDRKRTKAKSDAHSHTGSSHRGSPKHTETRHGRPTP	2783
Db	2157	LDSEVAHY-----GRTAAQAV-----SSGLASPL-----PREP	2184
QY	2784	ERSEMPSLHLGSPKKSASSSTRKAVSRN-----RSEIITYQLWMSGFKIGM	2827
Db	2185	PRARVRHVAEATGTEEDTVTLTAATEAVRVOLYELHIESNGLOGV	2231


```

Db 1017 -----EISRLQSKIKEMQ-----QATSPILMLOSGOVIGEVEBEDGALSTILOEQLEEE 1068
QY 1406 NMLK-----KEAEEYSKSEKNOFSLAEVPSGOKLVDEIVLKAOLKAABEELKID----- 1457
Db 1069 NGVULLSLQRAHBOAVKEN-----VKMAEISRLQOARLOKLEPIVMSCLDEP 1117
QY 1458 -RDYF-----ELVOTANTNLVEGKETETPLADHEED-----SIDRS---E 1494
Db 1118 ATEFGTARQOTQOFLQONTKQVEGVTNRHVLSLEDEEDVRLGISTGSSVQOKEKIE 1177
QY 1495 EMEIKVYG-EKLEBNOYLLERLOEKELESNKLEILOKEMETSVLLKDDLOKLESILSE 1553
Db 1178 ESBASVEGSELENS-----EETRTSEWELKNQISQLOEQMLMLCADCDDRASSEKKODLLFD 1233
QY 1554 NILLKENI-----DTTLKH-----HSDPTQAOLOKTOEQLOELANLALAADNCPITQOKE 1603
Db 1234 VSVLKKKMLKLERIPKASPYKLLYEEDVSRENDCLOELMRMETRYDEALEN-----NKE 1288
QY 1604 TSADCVHPLKEKILLITTEELHOKTNEOEKILHEKNELEQAQVE-----LKCEVH 1653
Db 1289 LTAH-----VRLQDELAKKMEVETETLSLEKSYDEKIEBGLNVLVRLQKIEK 1340
QY 1654 LKMSMIES-----KSSLESLOHEKHDTQOULLALQOQOVVTOEKELQOHEHILTAEV 1707
Db 1341 LOSVAVORCDDCLMEASLEMLEIE---PDGNITLONTLEBECVPRVSHVIEBCKOEN 1397
QY 1708 DHAKEINELGNKNEKNOQOTTKEOCLLNKKELOEQSOHLOCEI--E--LMSLKDK 1763
Db 1398 QYLEGNOL-----LEKVAHEIAWLGITQTHORPRVONVYLIEETVTLGFDK- 1449
QY 1764 SALETLESEKQVYNLQOEWMLMEELKNSORVIAERDOLODDLRSSEKSIETOP 1823
Db 1450 -----HQHQQTIAELIELEKTK-----LOELTRIKERVITLVQOKD 1486
QY 1824 DLKRAOEAALQOKVQVLTSLQISVLQEKISL-----ENOMLVNATVKEITLER 1874
Db 1487 VLSHGEKE-BELKAMMHDLQITSEMOKVYELLRYESEKLOEQENSILRNET--TLNEE 1542
QY 1875 DDLOSQKHFLSEIETLSLSKEKEFALBOAEDKADMAKKTIDITIKISINIEQULLQA 1934
Db 1543 DSIISNKK-----LGTINSQOEEWOMKTEFVKOENNAVQKVEMLKKQOIS--ELKIKINQ 1594
QY 1935 TNLKET-LYERESLQOKEOOLANTHELRITLKSKDALCKMEQREDEANKVYALTEKM 1993
Db 1595 LDEINTELSSKNS--QOKELOLQELNQLTLMLOKEKEPKN----- 1633
QY 1994 SSLQEOINENVTTLKEGEGKETFYIQR--SKOQSSQMEELRESLTKTKDLOLEBAEK 2050
Db 1634 SALEER-----DOEKNLKEKEILERCKVQSSSTLVSSLEALSEVKIQTHIVQ 1680
QY 2051 E-----ISEATNEIKNL--TAKISSLIEEILONASLINEAVSERENLRISKOLVSELOLS 2105
Db 1681 ENHLLDELKEMQOLHRCPLSDFOOKISSVLSYNEKLKEKELEKELSEBLSNCDVLKASS 1740
QY 2106 LTLKSRDHAFOKREKDELVNVIASIAEEIKITLTKEMDEFRDSKESLOSSHSELIC 2165
Db 1741 LL-----EHRATATKQOESKSEHOSASLSKQVAVSQEVOMLEDTIVONVLOMSMKDILR 1796
QY 2166 TYTTELQMLKQOKEEDINNKL--AEKYEVDELLOHLSLKEQLODQLOMELRNEKIRAYEL 2223
Db 1797 VTQOEKALQOEVMSLHKLOLQONAGKSMAPETIHPISGLHNO--OKRLSWPDLH--- 1849
QY 2224 CEKMDIMEKISVYLRMQNPEQOEDVABRMILLESRNOEIOELMKISAVVSEQHTL 2253
Db 1850 -----LMNEEQ-----LMOENRLOMTVMONVTAELTHSREKV 1883
QY 2284 SLSSELOKETEAKH--CMINIKESLSTLSRSGSLQTEHVLN--TOLQTLKLFK 2338
Db 1884 ROLESNLLPRKHOKHINSGMINPTEQEKLSLKRECOQFOKEQSANANKVQMSNLEBELE 1943
QY 2339 VYRTAAVKEHSHLIDYERKDLAEOKRHDELRLQLOCLEOHG-----RKWSDSASEEL 2392

```

```

Db 1944 TTH-----LENEGILKK-----KQVKLEDOLEMOHLKSTAPSPSPHAW-DLOLLQO 1989
QY 2393 KFCIEFLNELLFKKANIOSVODDFSEVOVPLNOVSTLOEELHKKGMOMLEEKGD 2452
Db 1990 QACWVPREQLOLQORLOLQ-----ERINQHOLELELN----- 2024
QY 2453 HVDAKKLESGMOENRRIASTIOLLTKRLKAVVOSKIOREITVYVNOFEAKLOEKKEQNK 2512
Db 2025 -TSETNTPQONQ-----QLVY-----VMEEMIEVEOKLKYKRLDQEKVQOK 2068
QY 2513 ELKRRMEHGBSASVMEENARLLGIITVODESKLQSRKMLNENL-----NLVKD 2565
Db 2069 EOICKNTRKADAMVADLYVENAOLKALEVTEQROKTAERKNVYLEEKRIASISNIVRN 2125

RESULT 11
US-09-727-384-6
; Sequence 6, Application US/09727384
; Patent No. US20020098511A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Helichman, Karen
; APPLICANT: Climbora, Daniel M.
; APPLICANT: Bush, Angie
; APPLICANT: Mauck, Kimberly
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 2318-271
; CURRENT APPLICATION NUMBER: US/09/727, 384
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,377
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/168,379
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/185,056
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 2139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-727-384-6

Query Match 5.4%; Score 794; DB 10; Length 2139;
Best Local Similarity 20.9%; Pred. No. 8.7e-23;
Matches 527; Conservative 444; Mismatches 912; Indels 634; Gaps 113;

QY 271 KKLSDGQAGF---INRDSKLR-----ILNSLGNKATVITICTTPVS 313
Db 21 KASDDVSGFHRKIQHVKNELCHMLSLAEVAPVLOQTLLQDNLGR-----VH 69
QY 314 FDELSTLQPASTAKHVRNTPHVNELDDEALL-----KRY-RKEILDKKOLENL 363
Db 70 FDFKEL-LLIISRTLSENEHFE--PDCSLEQPKYVGVGKGGRRSJPFEQSEVEEF 126
QY 364 -----ESSSEKAOAMAKEEH--TOLLAIKOLHNERERIRHNLNIVASSQES 411
Db 127 PEVTVIPLDEARPSIIPAGDSEHWKTO--RSEYEABGOLRFVNPDLNASSQSS 183
QY 412 -OODORVKKRRRVYAWAGKIONSILHAGVSDFDMLSRIPNFSKKAFFSMPSPFETIDS 470
Db 184 PQD-----WIEKIQEVEDGIGTRDGLNR-----KLVISICQYGLQNVGD 228
QY 471 VCFE---SPFDALSMND-----SNGIDAENKLAKVTHREKTSILHOSHIDFGQ---- 517
Db 229 MLEEVFNLPDGTMSVEDFEYGLFKNGSL--TPASVTPYRLKRLHLSMQSFDESGRR 285
QY 518 -----ISDSVQFDSKENVOLYLPKDSGDMACRKASFKEKTTSLQOQ--LOSKEBEK 570
Db 286 TTSSAMTSTIGFR-----VFSCLDDG--MGHASYERILIDQOEBEISQSLIK 333
QY 571 ELVOSFELKIALEBEOLSVAKAKNLEAVTNSREH-----SINAEVQDVER--EYVRKEMS 623

```

Db 334 ALDFSLDGNINLFE--LTJALENELLVTKNSHQALALASKAFIRHLEEDVDVAREK 391
 Qy 624 VLCDSGYNSNDLODSDVGKRLSSHDECIRHMKLEQIYDL--EPIENLNNKSEN 661
 Db 392 LRSDDLKAERKLSKLMASEVD--DHNAAIRRENYNLRKLDGEXKERIALKNELRK 445
 Qy 682 DKRSEODEMESIOCEALMAKNALEJALMRONFONILLENLTKREIALDERSLK 741
 Db 446 ERQIILQAGKORLEBOL--EKAATEEN--YIRDR-----LALSUK 484
 Qy 742 ENQETNEFELERKQEHBAOLIHGSLKLVENAMYNOMLEEDLETKRLKEOEI 801
 Db 485 ENSR-----LENE-----LENE-----LENE----- 498
 Qy 802 QJALEKRRADNLOKVRNPDLSYMGDSEKLCIEIFOLKOSLSDAAYVRDQKESFLR 861
 Db 499 KLAEYENLNRKLORNLEN--LAERFDDLPSSAEFLQOERLTO--MRNEYRQCRVLQ 554
 Qy 862 SENLELEKEMEDTSNMYNOKERASLEFEKOLETEKSNYKMEAD--LOKELOSAREIN 918
 Db 555 DOYDELQSELEF--YRAGRVL-----RLPLKNSPSEVEANSGSIPHEHLSGSECN 605
 Qy 919 YLNGLIAGV-----PRDILS--RVLEEKVSEFSQOLEKALEE--KNALEN--EYTCI 966
 Db 606 PLMMSIEAEIVIBOMKEOHHRDIOCLLEDEKVRHYEKOLDETVVSCKKAQENMKQRHE 665
 Qy 967 SEKKFJPNVEKELKNOISKASEIIMLKOGESASISKOETIMOQSOLOLDEYV 1026
 Db 666 NEKRTLEKOJSDUKNEIAELQOGAAYLK--EAHHEAT--CRHE--BEKOLQVLEBEKI 719
 Qy 1027 HQSKVOQTEEOYLEMKKMHDLFEKYIRNKSEADLLREMIKGTMSVE-----YK 1080
 Db 720 HIOEKRLQHE--MEIKARLTOQASFEREGLQSSAMTEKVRGULQOLEHQOELT 777
 Qy 1081 IADTKHELETTIDKROLHEKYPFOAMOTIPTIPPLSDSPSKLVBGSNDPIEIN- 1139
 Db 778 SLVEKHTLE-----KEEL--RKLELKHQ-----RELQEBERKEMTECNR 815
 Qy 1140 -----DYNNILALATERNNIMWC--LETENSLKEQYIDLNPLOLSLOAQISKSDI 1189
 Db 816 RNSQIENQOSDOQYTER-----CESALOSLEGKRYQOELKDQO-----OREKSO 864
 Qy 1190 QKPKOUL--EVEGKLLLEMLKGLJDSOLSIEKLOLENIETKLOLQOEMKNITI 1247
 Db 865 EFKEDLBTQECABAQOELKTKETKREKTSJVLTOEREMLE-----KTYKHLMSVY 916
 Qy 1248 ERRELOTNEDLKAHDSUKODISENIEOSIETODELRRA--QELRPOKOLVVSFOQL 1306
 Db 917 ERQOLLQODEDLKNNVSETOOSLSDQIL-----ELKSSHKRELREBEVL----- 961
 Qy 1307 DCSVGJISSP-----NHDAVANOEVK--LGEVNSLOSMLREBERDELQTSCK 1351
 Db 962 -COAGASEQOLASORLERLEMEHQ--ERQEMSKYLAMENHKTCTADREREMST--- 1016
 Qy 1352 ALVSELELRAHYKSVGENLEITTKLUNGLEK--ELIGKSE--ESEVLKSMU---ENIKED 1405
 Db 1017 ----EISRLQSKIKEMQ--QATSPLSMLQSGQVYIGEEVYGDALSLILOQEOELLE 1068
 Qy 1406 NMLK-----KQOAEYSSKRNQFSLSEVFSQOKIVDEIVLAKOLKAABERLETKD----- 1457
 Db 1069 NQDVLISLQRAHQAVKEN-----YKMAEISRLQORLOKLEPGLVMSGLDIP 1117
 Qy 1458 -RDYF-----ELVQANTNLVEGKLETPLOADHEED-----SIDRS--E 1494
 Db 1118 ALEFFPNTABQIBQLOQNRKQOVBGVTRRHYVLSLDEDEVRDUGSTGSSVQKQYKLE 1177
 Qy 1495 EMEIKVIG--EKTERNOYLLERLOEKELELNNKLEILOKEMETSVILKDKLOQKLESLSLSE 1553
 Db 1178 EESBASYEGSELENS-----EETRTESMLKNOJSQLOEOLMMLCACCDASSEKKODLDE 1233
 Qy 1554 NILKKNIT-----DTTKKH--HSDTOAQLOKTOOELOLAKNLAIAASDNCPTIOEKE 1603

Db 1234 VSVLKKKLMKLERIPASPRKYLLYEDVSRENDCLOEELRMETRYDALEN-----NKE 1288
 Qy 1604 TSADCVHPLERKILLTTEELHOQTNQOEKILHEKNLEQAOV-----LKCEVH 1653
 Db 1289 LTAE-----VRLQDEKKEEVEETFLSLEKSYDEKLENEGLNVLRLQOKIKR 1340
 Qy 1654 IMKSJTES-----KSSLESIOHEKHDTQOOLALAKQOMVYVQOEKELQOETHEHTAEV 1707
 Db 1341 LQESVQRODCCLMEASLENLEIE--PDGNILQNLQOTIEECYPRRSHVHIEBCKQEN 1397
 Qy 1708 DUKENIEJLGNFKNEAQOKTKTEOCLNENKLEDSQHRLOCEI--EE--LMSLKDE 1763
 Db 1398 QYLEGNTOL-----LEKVAHEIAMLHGTQTOHERPVRVQVYLEENTLILGPODK- 1449
 Qy 1764 SALETLKESQKVINLQOEMEMVLEMEELKNSQRFVIAERDLODDLEBSVMSIEROD 1823
 Db 1450 -----HFQOATTAIELELEK-----LOELTKLERATYILVKOD 1486
 Qy 1824 DLKRAQALQOQKQOELTSQISVLOEKISLL-----ENOMLYNVAATYKETSER 1874
 Db 1487 VLSHGEKE--BELKAMHNDIQITCEMOKVELLRYESEKLOQENSLRNEIT---TLNEE 1542
 Qy 1875 DDLNOKOHLFSIEITFLSLSEKKEPALQOAEKDKADAAKTDITIEKISNIEBOLLOA 1934
 Db 1543 DSISNLK-----LGTLNSQOEMQKTEYVQENNAVQKWEVNLKQIS--ELKIKNOQ 1594
 Qy 1935 TNLKET--LYERESLQCKROLANTHELETLKSKDLALGKMEQOERDEANKVIALTEK 1993
 Db 1595 LDLENTELQKNS--QONQKLOELNRLTEMLQCKKEGN----- 1633
 Qy 1994 SLEEQINENVTTLKEGEBKETFYLOP--SKOOSOMELRESLTKDLOLEBAK 2050
 Db 1634 SALEER-----EOEKFNLEKELERKQVOSTVLSLEAELEEVYIOTHIYQO 1680
 Qy 2051 E--ISEATNEIKNL--TAKISSLEEEIILQNASILNEAVSERENLHNSQOAYSELOIS 2105
 Db 1681 ENHLLDELEKMKQLOHRCPDLSFOQKISSVSYNEKLEKLEKALESEELNSCDKLAKSS 1740
 Qy 2106 LTLKSRDHAFAQSKREKDEAVNKTIASIEIKITIKEMDEFPSKESLOEQSHLSEELC 2165
 Db 1741 LL--EHNLATMKQKQKMEHQSAKLSQOLVANSQKQVONLEDTQVGNVLQMSRMSDLR 1796
 Qy 2166 TYKTELQMTKQOKEDINML--AEKYKAYDELLQHLUSLSEKLOLOQOMELRNEKLNTEL 2223
 Db 1797 VYIOEKKEALKQOEWYMSLKHQLOLNAGKSMAPELATHPSGLHNQ--OKRLSDMKLDH-- 1849
 Qy 2224 CEKMDIMEKEISVLRIMQNEPOOEDDVABRMDILSRSNOIOLEKISAVYSEQHTLL 2283
 Db 1850 -----LMBEBO-----LMOENBRLOTMQNTKALETHREKRY 1883
 Qy 2284 SLSSELOKETEBAKH--CMLNIKESLSLSSKSPSLOTEHVKLN--TOLQTLNKKR 2338
 Db 1884 RQESNULPRKHQHLNPSGTMNPTQEKLSLRECDQOFQKQSPARKVYSOMNSLQOELIE 1943
 Qy 2339 VYKRTAAVEDHSLIDYKDLAABOKRDELRLQLOCEBQNS-----RKSDSASEEL 2392
 Db 1944 TJH-----LENBGLKK-----KQVKLDEQIMQOHRSTATPSPSHAW--DIOLLOQ 1989
 Qy 2393 KICEIEFLMELFKRANITIOSVDDPSEVOVFLNOVSTLOEELBKQFMQMLEFGDL 2452
 Db 1990 QACPMVPRQOTLOLQOLLOA-----ERINQHLQOEBELNR----- 2024
 Qy 2453 HDYAKKLSEGMQOENRIASTIOLITRLKAVQOSKIQOREITVYULQOFAKLOEKKEQNR 2512
 Db 2025 -TSEINTPOGNO-----QLYV-----VMEBRMLEVQOKLKVYRRLQOEAQNOJK 2068
 Qy 2513 ELMRMEHHGPSASVMEENARLLGLTKTVODESKLKQSRIMKLENEI---NLVQD 2565
 Db 2069 EOLCKNTKADAMVQDLYVENAQILKALEVTJEOBKQTAERKKNVLEBKIASLSNIYVN 2125

RESULT 12
 US-09-820-843A-73

```

; Sequence 73, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEI
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: 063915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73
; LENGTH: 1805
; TYPE: PRF
; ORGANISM: M. genitalium
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: g11045905
US-09-820-843A-73

```

Query Match 5.2%; Score 761; DB 9; Length 1805;

Best Local Similarity 21.2%; Pred. No. 1.3e-21; Matches 437; Conservative 392; Mismatches 768; Indels 465; Gaps 86;

```

QY 678 KSENDKOKS-SEODFMESIQLCEAIMEKANALEMALRDNFDNI-----ILENETLK 730
DB 24 QSEYERKRSKTDVVKIEN-----QLKEIKSLDEELKNLKGKQAEDNPELD 71
QY 731 REIADERSLKENOETNEFEILEKETOKHEBAQIHEIGSLIKLV-----ENAEY 781
DB 72 KKHNLVVDL--NRIVNEYK--NFQFQKNHVDVSELDNTRFYKNELTRIQDENADF 127
QY 782 NQ-----NEEDLEFTKILKEQELQIOLAEKRRADLQKVRN-----EDLSVM 826
DB 128 NSKANTANFOANTHNKLNDFHLENOGTINRLNQTINQMILIDNVALLONPNTV 187
QY 827 GDSEKCEEIFOLKQSLSDAAYTRDAQKESFLRENTL--ELKEKEDTSNMYNOKKA 884
DB 188 EKNYVLLNIDVQLYNELDQLE-----NOKRLSTIEYENTYELVSADNELYENIDQN 242
QY 885 ASLPEKOLEBESYKMEADLOKELASAFNEIN-----YLNGLLAGVPRDL 933
DB 243 QIOEKHGYQYTRDELQELERQIQLTKQELVDKESALRYKIDDAFYINARLA----- 294
QY 934 SRVLEKVSFEKQOL--EKALKEKNA--LENEYTCLSEYK--FLPNEVECKLNQISKASE 988
DB 295 -----ELDDVAKQLSFQDGTIKONAHVEDKLVALNKEKORLNTQKEAFNLRQSAI 347
QY 989 EIMLLKQEGHSASIIIS--KOEIIQOESEOQLTLDVYTHQSKVQOOTEQYL-----EMK 1043
DB 348 DINKLOENELFAKHLEHQNEFQKQSDSLKLETEYKALQHKINERKNNSATKSEELL 407
QY 1044 KMHDDLEKTRKSEADLLREMNKLGTVESVETKIDTKHLEETIRKKEOLLHEKK 1103
DB 408 NOERLEPEK--RREDITLL-----TQASLEY--EHORRESSQLKDKQ--NEVK 449
QY 1104 YFQAMQITPITPLPSLSPSKLVEGNSQDPIEINDYHNLIALATERNNIMVLEPERN 1163
DB 450 QHQ-----NLEYAKKE-----LDKERN 467
QY 1164 SLAEQ-VIDINTOLQSLQASIEKSDLO-----KPKODEEGEVKLLLEMLLKGHLTD 1216
DB 468 LLDDQKQVDESAIFOLKEKVAQERKELELYLVKKQKQDQENE--LLFPEKQLQHOAD 525
QY 1217 SOLSIEKQLLENLEVTKLQ-----TLODEKMNITIERNELQTNEDLK-----AEDSL 1266
DB 526 FENLELEKQOELFEAKHALERSFTIKLEDKEDLWTKAQOIANEFSOLKTDKSKSADELM 585
QY 1267 KQDLSNEIOSIB-----TODELRAOELEROK-----QLVDSRPOOLD 1307
DB 586 LQNTYEMLOEKQKLFQERTYFFERNNAVLNRLQKREELLQKETTLDQLTKSEQDEL- 644
QY 1308 CSVGSISSPNH-DAVANOEKVSLGEVNSLOEML-RGERDELQTSCKAL--VSELELLRAH 1363

```

```

DB 645 -----INQREHKEVAVSEK-----QKEILGKLQDQFSQTSIANSKNLAEREMA--- 688
QY 1364 VKSVGEMLLETTKUNGLEKELELKSESEVYLSMLNLEKDNKKLKQAQAEYSKQNF 1423
DB 689 IKFKE-----KELEATEKQILDVNNAEVIOADLQNSLOERSELDONAKORLADF 741
QY 1424 -----SLEEVFSGSQKYVDELEVYKKAQKAEBERLEIKDQYFELVQNTANLVEGLE 1477
DB 742 HNDSLKLTNEYELSLQKLOELQTEAN-----OKOHSIONAYF-----EGELD 786
QY 1478 TPLQADHEEDSIDRSEMEIKVLEGL-ERNQYLLERLOEKL-----ELSNKL---EI 1528
DB 787 -KLNRKQAFNLNRKQYMEVDATKQRLSDRHQAL--NMQAELDRKTHELNNAFLNDA 843
QY 1529 LQKEMETSVLLKDLQKLE-----SLSENIILKENIDITTKHNSDTQAOLOK----- 1577
DB 844 DQKSLQDQLATVKEQOKLIDERSALKEQEFENAVGFRHRSNKTSQLQKLYELTK 903
QY 1578 -TQOELQAKMLATASDNCPTIOEKETSADCVHPLEKILLTEELHQ---KTNEOKL 1633
DB 904 QSEQOTKETELKTAFSQLOQDYOFELQKD-----QEFQIAKQRELDKL 950
QY 1634 LHEKNELEQAQVELKCEYEHLMKSMIESKSSLESLOHEKHDTQOQLLALQOQOVYTOEK 1693
DB 951 AEKNQYK-----LELDNRFQALQNGKQDTYQAQLELEREHQHLEQ 993
QY 1694 KELQOTHEHLTAEDVHLKENELGLNFKNEAQKTKKQCLNENKEL---EQSOHRLQ 1750
DB 994 TAFQANESLLKQREQLKQIA--FYHELK--RNQFALALGKRLFAKEQDOORQDQ 1047
QY 1751 EIEELMKSLDKESALETKSEBQVIVLNQEMEMVLMEMELKRSQRTVYAEBOQDD 1810
DB 1048 EIMNRFQEFKEYYDDEDAKRR-----ELELEKIRSL---SOSAVE 1087
QY 1811 LRESVMSIETQDDLRKAQALQOQKQVQELTSQISVLOERISLENNQLYNAVYKET 1870
DB 1088 LERKREKATDFTMLNKVQHTQINRD--QLNSQI-----ROFLERKNFORRS----- 1134
QY 1871 LSEBDLQNSQHLFSELET--SLSLKEKEFALQAKQKADAKRTIDTEKISNIEE 1928
DB 1135 ---NEANAKRAFLIKRLRSFASNLKQEKALATOKLEFDRDQOK----- 1177
QY 1929 QLLQOAT-NLEKTYLVEREST--IQCKEOL-ALNTEHLRETLKSDLATGKMOERDEAANK 1985
DB 1178 KELQOATIQLOQFQFKQONPQIEKQROVAIKTO--CEKLSDEKALNQLKVELKNSQT 1235
QY 1986 VIALTEKSSLEEOINENVTTLKEGEGEFTYLOPSPKQOS--SSOMELRESILTKDQ 2044
DB 1236 YLANKNKAEYQOQLOQKYTMLDKLENLETKDQLDKHRISIFARLTKFANDLRFKKQ 1295
QY 2045 LEPAEKLESEANTENKNTAIISSLEBEILLQNASILNVAVERENLRISKQYL----- 2058
DB 1296 LTKAQRYVDDKNRLKERNHLFSLNTERKRAVLEQOISYFQKQOATDALASHKE 1355
QY 2099 -----SELEQSLTKRSR---DHAFQSKREKDEAVAKIASLAEIKILTKENDEFDS 2149
DB 1356 VKKKEGELQKLLVELEFRKTYLNDPFAKFSQREFEQORLKLLELQKTLOTQNSNFK 1415
QY 2150 KESIQOQSSHLSEELCYTELQMLKQOKEDINNKLAKEYEVDLLOHLSLKEQLODI 2209
DB 1416 TKAIOETEN-----SYKRGMEELNFQKKEFD-----KNSRLVEYFPRKMDETERK 1461
QY 2210 QMELNRKLNRYELCEKMDIMEKISVYLRMQN-----PROEBEDVARMOLIESRN 2262
DB 1462 ESQV---KLVIKETOQRANLLEQAANKLIENTKNTIDREKELKAFKKQVDDDISTKQ 1518
QY 2263 QEIOELMEK---ISAVYSEQHTLLSSLELOKETEAKHGMINKKE---SLSSYTSRS 2315
DB 1519 KELNELLENKLLQOOSLIERBRAINSKDSLANKIETIKR-QLNIDKEKRYRLVDRKRLA 1577
QY 2316 FGSLOTEHVKLNTOL-----QTLNKKRVYRYTAAYVEDSHLINDYKDLAAEQKRDEL 2370

```

Db 1578 EOKYQTEINRLTQTFDESEKODIKNEPPLFKI-----NGNDMAFPYLWLPQOKODN 1633
QY 2371 RLQLOCEJGRKMSDASSELKCELEEFINEL--LEFRANIIQSODDESEVQVPLNOV 2428
Db 1634 TLQRLQLEF-----QOLQPMOORYENELNEIRLROHNLLEKLD--QIO----- 1674
QY 2429 GSTLOEELERKGFOMWLEFEGDLHYDAKILSGMOENRIASTIQLTRKKA--VVO 2486
Db 1675 ---LESQLNKKQ-----SEFSKVESMEKLEKTESRLNDPQOKINYLTKKYNQNHNTYQ 1725
QY 2487 SKQRETTYLVN-----QFEAKLOEKKQNKELMRHMHGPPSASVMEENARLLGLIKTV 2542
Db 1726 PSSYQPPSPYSDSKQQLLFRIOLEKRON--LEQOQFOPAP--AVVOQPSFAAPNITK-- 1780
QY 2543 ODESKKLOSRIKMLNELNLVK 2564
Db 1781 -----QOOIADLMAEINNIK 1795

RESULT 13
US-09-885-535-4
; Sequence 4, Application US/09885535
; Patent No. US20020104105A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Helichman, Karen
; APPLICANT: Baritel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 2318-266-II
; CURRENT APPLICATION NUMBER: US/09/885,535
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213, 245
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2835
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-885-535-4

Query Match 4.9%; Score 720.5; DB 10; Length 2835;
Best Local Similarity 18.8%; Pred. No. 7.3e-20;
Matches 515; Conservative 508; Mismatches 954; Indels 761; Gaps 111;

QY 315 DETLSTLOPASTAKHVRNTPHNEVLDDALLKRYKEIILDKOLENLESSESTKAQAM 374
Db 117 DNKLSSTLAVST-----HPDAMNQOLETAQKMKQOEIQQEKKQIKVAQALCE-DLSAL 167
QY 375 AKREHQ-----LAEIKQLHREDRIWHLTNIVASSQESQODQRYKRRRTW 425
Db 168 VKEEYLKAEISRLQELILSKFDVEQKAEHNVHQLQS-ACASSHQPQMSRDFQ---AM 222
QY 426 ARGK--IQNSLHSAVSDFDMLSRLPGNFSKAKAFSDMPSPEID--DSVCTEFESQFDA 481
Db 223 LQTKREQKSH-----PISAKDLVLESLLIDHDKPFSST 256
QY 482 L---SMADNGSIDAEMNLASKYTHREKTSIHQSMIDFGQISDSVDFHDSKENOLQYLPK 538
Db 257 LTAQSHMYEKTIAEGENLLFKTQSGSEKALQLOLNTIKTNMPTFNQVAKERENKL----- 311
QY 539 DSGDMAECKKASPEKITLQO-----QLOSKKEEKKELVOSFE-----LKTAELE 584
Db 312 -----KESLEKALKYKEQVETLMPWIDKQNNLEIKELCDPAESENSIAKLKSLQ 362
QY 585 EQLSYAKNLKLEWYNSREH-----SINAEVOTDVEKEVYRKEMSVLGDGYNANSNDLQD 639
Db 363 KEMDQHFQGNVELLNNANSLVCELDKRYVTDENKSLQK-----VDWYT 408
QY 640 SSVDGRKLSHDECTEHRKMLEQKIVDLEEFIEMLNKKSEMDKQSSQO--DFMESIQLC 698
Db 409 EQLHSKKF-----CLEN--MTQKFERQE-----VSKESKRLQCAKAKQQLDIDHSL-- 452

QY 699 EAIMAKKANALEELALMRQNFNIILENETIKREIADLERLSUKENQETNEFELEKETOK 758
Db 453 -GSQATSNKTYITMLQTOQSLQALKHQVDAKLADLVASDSGTSDV--LLOVETIA 510
QY 759 EHRQALIHIEIGSLKLVEN-----AEYQNQLEEDLETFTKLLKLEQEIOLAEI 806
Db 511 QEHSTLSQVQDEKCSPLFETRIQIGHFQNTIRMFQSFAPFEDLDS-----MAYV 561
QY 807 KRKADLQK--VRNF-----DLSVMDSEKLCETIQLOKSISDAEAVTRDAQ--K 855
Db 562 GRDAETLQOKETIKAFLEKLEALMASNDNANTCKMMLATEBSPDLVIGIRDEALSK 621
QY 856 ECSFLSEMLEKEMEDT-----SNWYNQKRAASLFQK-----LETEKSN-- 898
Db 622 QCNKLLDRQARQEQVEGITKRLKEEYYSKLESILLQAAEHESQSGVGHETETINQ 681
QY 899 ---YKMEADLOKELOSAFNEINYL--NGLLAGKVPRLDLSRYE-----LEKKV 942
Db 682 LNMFKVFOKEEIEPLQGGQODVNMIGQGLIQSNAKSTQGLEHLDLDVNAHMKTLNKKV 741
QY 943 SEFSKOLEKAL-----EKNALNEVTCSEYKFLPNEVECLNQSISKEEIMLKQEGE 998
Db 742 AQRAAQLOELALHCGRFQDALE-----SLSMWDTIELVANO--KPSAEKRYK-- 790
QY 999 HSASIIKQEIIMQESQEIQLDTVEVHTQSQVOQTEQOYLEMKMHDLFEKYIRNKS 1058
Db 791 -----AQIQGKL-LQRLDD-----RRS 808
QY 1059 EAEDLIREENLKGTYESE-YKI-----ADTKHELEETIRDXEQLH 1100
Db 809 TWEVIRKEGEKIATTAEPDAKVKILQSLDSRWEALLNKETRRLQEGISVAAQDFH 868
QY 1101 EKKYFQAMQITFPTPLSDSLPPSKLYGNSQDPLEINDNLALATERNNIMVC--L 1158
Db 869 E-----TLEP-----LNEW-----LTIETKRLVNCPEI 891
QY 1159 EFERNSLKEQVIDLTQLOSLQASQIESKDLQPKODLEEGEVKLLLEMLKGLHLDQ 1218
Db 892 GTQASKLEEQIAQ-----HKVLQEDILLRKQNVQDA--LNLGIELIKQTTGDEV 938
QY 1219 LSIENKOLENLEVTETKQLQLOEMKNITIERNELQTNEDLKAHDSLKODISENIOSI 1278
Db 939 LTIQ-----DKLEIKARYDIT-----KISTDAKTEAL 970
QY 1279 ETQDELRAAQEELREQKQVLD-----SFRQQL--DCSVGISPPNH--DAVANOEFVSGE 1330
Db 971 QIARLRHSHHELCWMDQVEVELLSYEQVYLKGEASQAGRPRLKEKKAANNKALLDS 1030
QY 1331 VNSLOSMLRGERDELQJSCALVSELELLRAH--VKSVEGENLE-----ITKRLNG 1380
Db 1031 LNEVSSALL-----ELVPWRAREGLEKMKVADNDRYRLVSPDTIQKVEE 1074
QY 1381 LEKEILGKSESEVLSKMLENKEDNNKL-----KEQAEYSKSENGFSLE-- 1426
Db 1075 IDAALIRSQQFQDAADAEISWTETETKLMISGDIRLDDQTSQAOLOVQKTFTEILRHK 1134
QY 1427 ---EVSQSOKLV-----DELEVLKAQ--LKAARELEIKIDRDYELVQNTANLVE 1473
Db 1135 DIIDDLVNSGKHIMACSEBKQSKKLDVKLVNYDTIQNSRYQLBERA--QSLVN 1192
QY 1474 GKLEF-----PLQADHEDSIDRSEMEIKYLGKLEBNQYLLERLOEK----- 1519
Db 1193 QWETYEELMPWLTETQSIISQLPAPALEYETLRQOOEHHQLRELLIAHHPHIDKMNKT 1252
QY 1520 ---LELS--NKLETLQKEMETSVL--LKQDLQK--LESLISENIIKENID-- 1562
Db 1253 GPQILELSPGBEFSIQEKTYVAADTIYSQIKEDYKRAVALDAISQSQFQHDKIDQILLES 1312
QY 1563 ---TLKHSDDTQAOLOKTOQLOLAKNLALASANDCPTQO-----E 1601
Db 1313 LERIVERLRQPPSISAEVETIKQESKKNVSVDEKIKQPIYETPLKQNGEEMIAKSGTD 1372
QY 1602 KETSADCVHPLBEKILLTLTEBHQKTNEQEKILHEKNLEBOAOVELKCEVHLMKSMIES 1661

```

Db 1373 KDLSAKAVOKLDMVFIWEIHIHLVEERAKLLDVMEAE---KFMCDHSLVITIKDT 1429
Qy 1662 KSLSESLQHEKHDTQOQLLAKQOMVYTOEKKELQOQTHEHLAEDVHLKENIELGLNFK 1721
Db 1430 QDFIRLEDEPGIDPS---VVKQO-----QEAETIRREIDGLQELDIVINLG 1474
Qy 1722 NEMOQKTKKQCLLNKEKLEQSOHRLOCEIEELMKSJKOKESALFETLKESEQVINLNO 1781
Db 1475 SE-----LIAAGBPDPYKKSIDELNSAMDSLNKAMKDR---IDKLEEMQAAVOYOD 1526
Qy 1782 EMEVMELEMEELNSQRTVIAERDLODLRESEVMSIETODDLRKQOALQOQKQVOE 1841
Db 1527 GIGAV-EDWDIANGK---LASMPIGTDL-ETVKKQIE---ELKQFKESEVQOQIEMER 1578
Qy 1842 LFSQISVLOERIS-----LLENQMLVYAVYKETSERDILNOSK-----1881
Db 1579 LNHQAEILLKKVTEESDKHTVODELMELKLIV-----DSLERRITINROHKBGALLALG 1632
Qy 1882 --OHLFSEI-----ETLSLSLKEK-----EFALBEAKDKADAARTIDITEKI 1923
Db 1633 QFOHALDELAMLTHTGLLSEQKPVGGDPKRAIETELAKHHVLDONDVLAHOSIVEAVNKA 1692
Qy 1924 SN--IEBQLQOATNLKETIYERESLIQCKEQALANTEHLEETL-----KSKD 1969
Db 1693 GNDLIESSAGEEASNLQNKUL---EVLNQRQNVLEKTEQRKQOQDGLARQAKGFHEIED 1749
Qy 1970 LAIGKMOEODEAANKYI-ALTEKMSSLEQINENVTTLKEBSEKETFYIQRPSKOOS 2028
Db 1750 LOQMLDTEHRLHLASKPLGLPE---TAKEDOLNMEVCAFAKEKETIY---KSLMQKG 1802
Qy 2029 SOMELERESLKTDLQLEAEKEISEADNEIKNLTAKISSLEETILONASILTNVASEBRE 2088
Db 1803 QOM--LACRKSATNTDQ-----DINNLEKEMVESYETLNEKTKLEBALNLAM 1850
Qy 2089 NLRHSKOALVSELEQSLITLK--SRDAFAQSKREKDEAVNKISLABEIKITLKEDEEF 2146
Db 1851 EFHNSLODFINMLTQAOQTLNVAASPSLI-----LDTVLQFQDEHKVFANEVNSH 1900
Qy 2147 RDSKESLQEOBSHL-----SEELCTYTELOMLKQOKEDINNKLAEKVDELLOHL-- 2199
Db 1901 REQIETLDTKGTHTLKYFSOKQDVLINLNLSIOSRWEKVVQRLVERRSIDARAKRAQ 1960
Qy 2200 -----SILKQOLDQLOELNREKLRNYELCEKMOIMEKEISVLMQNEPOQOEDDAER 2254
Db 1961 FHEKMSLMELESESLSE---LEIANDPKIKITOLA-----QKFEPOKSLGAK 2009
Qy 2255 MDLIESHNOETLOELMERTISAVYSEQHTLLSLSELOKETEAHRCMLNTIKESLSLTSR 2314
Db 2010 HSYVDITNRTGRSLKEKTS--LADDNLKLDLMDLSLRDKMOT--ICGKSVROKMLEBAL 2065
Qy 2315 SFGSLQTEHAKLNTQLOTLNKFVYVRYTAAVKEDHSIKDYBKDLAEOQRHDELRQOL 2374
Db 2066 LFGSQFTD-----ALQALD--WLYRV-----EPOLADEQVPHGDIDILVM 2103
Qy 2375 QCLEOHGRKMSDSASEELKECEIFELNELFKKANIIOQVQDDPSEVOVFLNQVSTLOE 2434
Db 2104 NLIDNH-----KAFQK 2114
Qy 2435 ELEHKKGFQMLLEEGDLHVDAKKLSRMOENRRIASTIOLLTKRLKAVVOSKIQREIT 2494
Db 2115 ELGKRTSSVQALKR-----SARELLIGSRDSSVMVQOMELSTRETYCALSISKO-- 2166
Qy 2495 VYLNOFAKIOEKKQOK-----ELMRBHEHGPSASVMEERARLGLITKQOD 2544
Db 2167 ---TRLPALRQAEFSESVVHALLEMLAEQTLRFHG---VLPDDDEA---LRTLID 2215
Qy 2545 ESKLQSRIKLENELNLVDDAMHKEKVALILODKLISRAEALNMAQVYKLTQKODNL 2604
Db 2216 QHKFEMKLEKRAELN-----KATMGDYVALICHPSITTIKIMIIILIRARF 2264
Qy 2605 QAAKKEIENLOKMYAK---GAVPYKEEIDNL-----KITVVKIEMEKIKYSKATDOEI 2654

```

```

Db 2265 EEVLAAMAKHOORLASALAGLIAKOLLEALLAMLQNAETTLTDKQKEVI-----POEI 2318
Qy 2655 AYLSCELDREKREGLRLKEELRRQADNDTVCYKPKDQKASTFPYVCGGSGGIYSTAM 2714
Db 2319 EEWALIAEHQ-----TEMEENTRKQOPDYK---YTKYTKRRAAPSS-----2358
Qy 2715 LVLOSEKALERELSHYKKKKYHHLSTRWSSSEDRKTK 2752
Db 2359 --LOSHIPVLDKGRAG--RKRPASSLYPSGSQTOIETK 2393

RESULT 14
US-09-927-597-4
: Sequence 4, Application US/09927597
: Publication No. US20030032018A1
: GENERAL INFORMATION:
: APPLICANT: Malik, Fady
: APPLICANT: Berard, Christophe
: APPLICANT: Freedman, Richard
: APPLICANT: Craven, Andrew
: APPLICANT: Sakowicz, Roman
: TITLE OF INVENTION: Human smooth muscle myosin heavy chain
: FILE REFERENCE: CYTOPO18
: CURRENT APPLICATION NUMBER: US/09/927,597
: CURRENT FILING DATE: 2001-08-10
: NUMBER OF SEQ. ID NOS: 17
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 1979
: TYPE: PRT
: ORGANISM: Human
US-09-927-597-4

Query Match
: 4.98; Score 719.5; DB 9; Length 1979;
Best Local Similarity 19.88; Pred. No. 5.5e-20;
Matches 462; Conservative 428; Mismatches 847; Indels 591; Gaps 93;

Qy 17 IOROGDQANLQW-----KAGNNTISOVDGTSFNFDRVFNSHESHSOIYOELAVPII 69
Db 50 IKKEGDEVVVELVENGKKTATVGKDDIOKNPK-----FSKYEDMAELTCLNEASVL 102
Qy 70 RSALQGY-NGTIFAYGQTSQKTYTMGTSPNSLGIPOAIOEVF--KIIQEIPIREFLR 126
Db 103 HNLRERYFSGLIYY---SGLFCVYVNPYKHLPIYSKIDYMKGKRHHMP-----151
Qy 127 VSYMEIYVETVKLLCDDRKKKPLEIFEDENRNVYV---ADLTEELVMPEN--VYQW 179
Db 152 --PHIY-----AIADYVRSMLQDRD--QSILCTGESGAGTENTKKYIOYLAVVAS 200
Qy 180 IKKGKRNH-----YGE-----TKMNBHSSRSHTIFRMIVESR 212
Db 201 SHKGGKIDSTIGSPFAYGELKEKOLLANPILAEFNAKTKYKNDSSRFGKFIK-----255
Qy 213 DRNDPTNSCND--GAVVYSHLNLVLDLAGSERASQGAEGVRLKEGCNINSLFLGVI 270
Db 256 -----NFDVGYIYGANIEYILEKSRAIRQARDE-----RPFHIEYMI 295
Qy 271 -----KRLSDQAGGFIVY-----RQSKLRIILQNSIGNAKYIITIPV 312
Db 296 AGAERKMSDILLGFFNNYTELNSGEVPIPAADDEMFO-----ETVEAMALINGF 345
Qy 313 SFDETLSTLOFASTA-----KHVRTPPVNEVLDEALLK-----RYRKEIL 354
Db 346 SEEBOLSLTKVYSSVLOLGNIVFKKERITDAS--MPDITAOKYCHLMGINVTDFTRSL 404
Qy 355 DLK--KOLENLESSSETKAQ-----AMAKEBHTOLAEIKOLHKEREDRIWHLTNIVAS 407
Db 405 TPRIKVGDDVYQAKQTRQADFAVEALAKATYERLER-----WILTRV-----447
Qy 408 SQESQODQVRKRRRYTAPKQIONSILASGVSDPMDLSRLPGNFSKAKKSDMPSP-- 465
Db 448 -----NKALDKTHROGASPLGIL-----DIAGEEIF 473

```

QY 466 EID--DSVCTEESD-----FDDALSMADS-----NGIDAEMNLAKVTHREKTSLHQM 512
 Db 474 EVMSFEOCLINTYNEKIQOJFNHTMFLBEOBEVOREGI--EWN-----F 515
 QY 513 IDRG-QIDSVDQF--HDSKENDLOVLPKDSGDMACRKA SPEKEITSLQOOLQOSKEEK 569
 Db 516 IDGLDLOPCIELIERENPNPGVLLALDECEWPFKATDKSFVEKICT--EQSGHPKQPK 573
 QY 570 KEIYOSPELTAIELEBOLSKA-----KLEMTYNSREHSINA---EVQDVKEVYRKE 621
 Db 574 KOLKDKTESITIHAGVVDYNASAMLTCKNMDPLNDVNTSLINASSDGFVADLWMDVDR-- 631
 QY 622 MSYLG--DSGYNASNSDLODSVDGKRLSSHDECIEHRKMLEOKIYDLFEFIENTLAK--KS 679
 Db 632 --IVGLDQAMAKMTESSLPASAKTKKGMFRTVGOL--YKEDLGKIMTLTLMRTTPFVACIT 687
 QY 680 ENDKQSSSEOD-FWESTQL-CEALMAEKANALELALMDRNPDIILENETLKR-ETA-- 734
 Db 688 PNHKRSKGIDAFVLEQLRCNGV-----LEGIRICRGFPRIYVQFQFROREYILAA 740
 QY 735 -DLERSLKEQNEFEI-----LEKETQKHEAQLHEISLKLKLVENAM-----Y 781
 Db 741 NALPKGFMDKQACITMLKALELDPNLYRIGQSKIFPRTGVLAHLEBERDLKITDVIMAF 800
 QY 782 NQMLEDLKTKLKEOEIQLAELRRADNLQKVRNPD-----LSYMGDSE 830
 Db 801 QAMCQVLAARAKAKROOJLAKVIOIRNCAVILKLNQMWMLRTFKVRLLOYTRREE 860
 QY 831 KLC--BEIFOLKQSLDAEAVTNDQKESFLSENLEKKE--DTSNNYNOKERAST 887
 Db 861 MQAKRELOKTKEROQAEHELKELEQKHSQLEEKMLLOEOLOAELELAEEAEEMVRL 920
 QY 888 FEKOLETEKSNYKMEADLKEIQLASAPNEINYLGLAGVPHDLRSVLE--KVYSEFS 946
 Db 921 AAKQELIELILH--EMEARLEE-----BRCQOLOAERKKAQOM 959
 QY 947 KOLEKALEEKNALE-----NEVTCLESEKFLPNEVECLNQISKASEIMILKOEGBHSA 1001
 Db 960 LDLEOELEBEAAROKLOLEKVTAAEAKIKLEBILVMDQNNKLSKERLLEER----- 1014
 QY 1002 SITSKOEIIMOEOSELOLTDVETHQOSVQOOTEBOYLKMKMHDDLFKRYRNSSEAL 1061
 Db 1015 --ISDITTNALAEERKAKNTLTKLNKHSMSILE--VRLKREE-----KSRQELE 1061
 QY 1062 DLIREMENLKGTESEYVKAIDTKEHELEETIRKEOHLHEKKYFQOMQOTIPPTPSDS 1121
 Db 1062 KLRKLE--GDASDPHEQIAD-----LOAQIAELKQOLAKKEBELQA-----ALARLDE 1109
 QY 1122 LPPSKLVEGNSQDPIEINDYHNLALATERNNIMVCLTERNSLKEQVIDLNTQLOSLQA 1181
 Db 1110 I-----AKNNML-----KRIRELEGHISDLOEDLDSERA 1139
 QY 1182 QSTIEKSDLOPKODLEGEVKKLLEMLKLGHLTDSOLSTIEKLOLENTLEVEKLOIOT 1241
 Db 1140 ---ARKNAEKQKRDLCG-----ELEALTELEEDT-----LDSTATQOELRAKREQ 1181
 QY 1242 MKNITTEENELQTFNEDEKKAHDSLKODISENTEOSTIETODELRAQOELREKOLVDSE 1301
 Db 1182 --EVTYUAKRLDEBETSHENOVOEMROKNAOAVBELTEQOFRKRAKANDKKNQOTIEKE 1239
 QY 1302 ROOLDCOSVGISSPNHDVANAOKESVLGEVNSLOSEMLRGE--RDELQTSCKALVSELEL 1359
 Db 1240 NADLAG-ELRVLGQAQOEYVHKKKKLEAOVQEOLOSKOSDERARAELENDVHKLQNEVES 1298
 QY 1360 LRAHVAVSGENLEITIKKNGLEKEKELIGKSESEVLSKSMENIKEDNNKIKEOAEYSSK 1419
 Db 1299 VTEMLNEABEKAIKIKADVASSOL---ODTOELLQOETFRKLVNSTKIROLEEBENSL 1355
 QY 1420 ENQFSLEEVYSSQKLVDELEVLAQOLKAAEERLEIKDRYFELVYQATANNVLVEKLTLP 1479
 Db 1356 QDO--LDEMEKAKONLEHRISTINLOLSDSKKLLQ---DEFASTVEA---LEEGK--- 1401

QY 1480 LOADHEEDSIDRSSEMEIKVLGEKLERNOYLERLOEKELELSNKLEITLOKEMETSVLL 1539
 Db 1402 -----KRFQEIENLNGOYEERKAAVADKLEKTNRNLOQELDDLVYLDNQRL 1449
 QY 1540 KDDL---QOKLESLSLNTILKENIDITTLKHSDYQAOLOKTOEOQLAKNLAIAASDNC 1596
 Db 1450 VSNLEKQORIKDOLLAE---EKNISSKYADEBD---RAEAE----- 1484
 QY 1597 PIIOEKTSADCVHPLEEKILLTLEELHOKTNOEKLHEKNLEBOAQVEKCEVEHIMK 1656
 Db 1485 --AREKET-----KALSTARAL-----EELKEKLEERKNKMLKAMEDVYS 1525
 QY 1657 SMIESKSSLELOHEKHPDQOLALQOMQOYVTOEKKELQOTHEHILTAEVHILKENIEL 1716
 Db 1526 SKDQVGNVHLEKRSKALETQMEKMTQOEELEDELOATEDAKRLLEVYMQALKGOFER 1585
 QY 1717 GLNFKNQOQTKTQOCLNKENKELESORLOCEIPELMK-----SLKDK 1762
 Db 1586 DLQARDEONE-----KROLOROLHEVTELEDEKROBALAAAKKLEGLDKDL 1636
 QY 1763 E---SALETKSEQKVINLQEMEMVMELEIKNSORTVIAERDOLQDDLRSEVMS 1818
 Db 1637 ELQADSAIKGREERAIKQRLQOMKDFORELEDARASROEITATYAKENKAKKSLEADL 1696
 QY 1819 IETQDDLKKAQBALQOQKQVQELTQISVLOEKISLLENQMLYNATVYKETLSERDGLN 1878
 Db 1697 MQOEDLAALAEARAKQADLEKELEAELA-----SLSGRNAQ 1735
 QY 1879 QSKOHLFSELETSLSKKEFPALQAEKADAKRTIDITEKISN----- 1925
 Db 1736 DEKRRLREARIAQLEELLEEQGNMEA---MSDRVRATQOAOULSNEALATERSTAOKNE 1791
 QY 1926 -IEOOLQOATNLKETLYEESLQOCEQLALNTEHLRETKSKDQALGMEQORDEAN 1984
 Db 1792 SARQOLERQKKEKLSKHEGAVKSK-----FKSTIALLEAKIAQLEQVQDEAR 1842
 QY 1985 KVIALTERKMSLEQJINENVTTLKEGGEKETEYLORPSKQOSSQWELRESLAKTKDIO 2044
 Db 1843 EKOAAFTSLKQKQKILKEILLQVED-----EKMAEQYKQOAEKGNANAKVQLKQ 1892
 QY 2045 LEEAEKESATNEIKNTAKISLLEELLQANSIINENAVSEENLRHSQOQVSELEOL 2104
 Db 1893 LEEAE-----ESORINNRRLQORE-IDEATESNEAGREVNALKSILR----- 1936
 QY 2105 SLTKSRDHAFAQSKREKDEAVNKIASIAEIKILTKEMDEFRDSKES 2152
 Db 1937 ----RGNETSFPVSRSGRVIENMDGSEF-ETDTRDAD-FNGTKAS 1978
 RESULT 15
 US-09-815-242-12996
 ; Sequence 12996, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12996
LENGTH: 6281
TYPE: PR
ORGANISM: Staphylococcus aureus
US-09-815-242-12996

Query Match 4.98; Score 718; DB 10; Length 6281;
Best Local Similarly 18.56; Pred. No. 2e-19;
Matches 623; Conservative 617; Mismatches 1277; Indels 854; Gaps 150;

1 MSEDVAVCVVRPLIOREGDQA-NLQWKAANNISQVGTSGFNFDRVNSHESQ 59
DB 3054 IADDAKATAYNFTDADAKO--QAYNTAVTNAENIISKANG-----GNATQ 3098
QY 60 IYOEIAVPIIRSAQYNGTIFAYGOTSCKTYTMTGTPNSIIGIIPQAIQEVFKIQIEIP 119
DB 3099 AEVQAIKQVNAKQALNGN--ANVQAKKDEATALINSNDLN--QAKQALK--QYVQ 3151
QY 120 NREPLKAVSIWEIYNEVKDL-----LCDDRKK-----PLEIREDPNNRY 161
DB 3152 NATVAVAGVNNV--QTAQELNNAMTOLKQIADKEQTKADGNFVNADPDKONAYNAVA 3208
QY 162 VAD-----LTRELVAWPEHVQIWK--GEKNRHYGETKM--NDHSSRSHITRMIVESDR 214
DB 3209 KAEALISATIPVVTPEITRALNKVTOAKNDLNGNTNLATAKQVQALIDQLEPILNQAO 3268
QY 215 NDTPTNSENCDGAVVSHNLVDLAGSERASQGAEGVRLKEGCNINSLFTILQVYIKLIS 274
DB 3269 RD-EYSQIQTQATLVPNVNAIQOA--ATLNDAMTQVLSKGI-----AN 3308
QY 275 DGAQGGITNRDCKLTRLQNSLGGNAKTVIICITTPVSFDETLSTLOFASTAKHVRNP 334
DB 3309 KAOIKGSENYHD-----ADTDR-----QTA 3328
QY 335 HVNEVLDEALKRYRKEIIDLKOLENLESSSEFKQAQMAKEET-----OLIAEIK 387
DB 3329 YDNAVTAAEELL-----KOTNPYDPTIQAALIKVNDTQNALNGNOKLADAK 3377
QY 388 QLHKEREDRIWHLTNIVVASSQESQODORVKKRRVTWAPKIQNSLHAGVSDPDMLSR 447
DB 3378 QDAKTITGLDHLND-----AQKQALFTQVEQAPDIATVNNVKNQAOQLN-----NMTN 3427
QY 448 LPQFESKAKAFSDMPFPEIDDSCTEPPDDALSMDSGIDAENNLASKV--THREK 505
DB 3428 LNNALQDKT-----ETLNSI--NETDADAKKDAYTAAVSHAEILISKANGSNASQ 3476
QY 506 TSLHQSIMIDGQISDVQVHDS--SKENQLOVLPKDSGDMAEGRKASFKEITSLQOOL 562
DB 3477 TEVEQAMORVNEAKQALNGNDNVORAKDAKQV--TNANDLNQAKDKLQOVDA--AQT 3533
QY 563 QSKEEKKEELVQSEFLIAELEQOLSYKAKNLEMTNSRSHINAEV--QTDVEKEVYRK 620
DB 3534 VANNVTIKQTAQDNLQAMTQKQIADKQ-----TKANGFVAADDIDKQANAYNAVAHA 3588
QY 621 EMSVLGDSGYNASND-----LQDSSVDGKRILSSHDECIEHRKMLEQKIVDLEETENL 676
DB 3589 EQLISGTPNANVDPQVQAALQVNAQKGLDNGNH-----NLQYAKDVAN 3633
QY 677 KKSE-----NDKQSSRQDPMESIQCEALIAEKAANALELALRDNFNILLENITLKR 731
DB 3634 TAIQDLNLNMQPQKTAALKDQVSHAEVLTVGNALIKQNA-----DALNNM--GTLKQ 3662
QY 732 EI-----ADLERSLKENQETNEFE-----ILEKTOKEHQAQLIHEI 768

DB 3683 QIQANSQVPSVDFTQADQDQQAAYNNAQAQALANGIPPIVPLPDI-----VTQAVTMT 3738
QY 769 GSLKVLVNAEKYNNONLEE--DETQTKLKEQ-----EIQALERKADNLOKRYN 819
DB 3739 NOAKKALNGDEKLQAKQALANLDTLRDLPQCDALRNOINQAOALATATEQTKQONQ 3798
QY 820 FDLVSMGDSKELCEITQIKQOSLSDAEAY-----TRDAQKESFLRSNLEKEME 872
DB 3799 VNTAMS-----NLQGIANKDTVKASENYHDADDKQVYNAVSAQEGITNQ 3846
QY 873 DTSWYNQKKAASLFEKQLETEKSNYKRMKMEADLOKELQSAFNEINLGL-----LA 925
DB 3847 TTNPTLNPDEITRALTO--VTDAKNGLNGEAKLATEKQONKADVSGHTLINDAQKALK 3903
QY 926 GKVPKDLISRYELKEKVEFSKOLEKALEE--KNALENEVTCISEKFL--PNEVECLKN 981
DB 3904 GQI--DQSPETATVAVQVQATISLDAQMDQSLQALNDQAQTLADGNYLNADPDKONAYK 3961
QY 982 QISKASEIMLLKQEGHSAIISKQETIMQESQIQLDDEVYHTQSKV-----1032
DB 3962 AVAKA--EALINKQSGTN-----EVAQAVESTINEVNAKQALNGNDNLNA 4006
QY 1033 -QOTEEOYLEKKKMD--DLFEKYIRNKSEAEDL--REMENLKGTMESEVYKIAOTK 1085
DB 4007 KOQAQOLANLTHLNDQAKQSPESQITQAPLTVDTVTINQKQTLDHAMELLRNSVADNQ 4066
QY 1086 HEL-BETTRDKQQLHEKKYFFQAMQTFPTPLPSDSLPSKLVGNSQDPLEINDYHL 1144
DB 4067 TTLASEDHD--ATAQRNDYNOAVTAANNI--INQTSF-----TWNDDVAV-- 4110
QY 1145 IALATERNNIMVCEETERN--SLKEQV--IDL-----NTQLOSLQASIEKSDP--QK 1191
DB 4111 -GATTVQVNTKVALDGDENLAAKQANRRDLQDLHNNQAKQLOQOSITQSSDIAAVNG 4169
QY 1192 PKODLEEYKILLMEMLIKGHL-----TDSQLS-----TEKLOLEN 1228
DB 4170 HKQTAESLNTAMGNLINALHQAQVEGRGFINDQTKQATVYNAVNAAMINQOTQON 4229
QY 1229 LEVTEKLOTLOEKMNITIERNELOTFNEEDLKAEDSLKQDSENIESTIEF-----1280
DB 4230 ANQTEVEQAI-----TKVQTTLOALNGDHN--LOVAKTNATQALDALSLNDPQ 4276
QY 1281 -ODELRAA-----OELREKQOLVSPFOQLLDS-----VGISSP--N 1316
DB 4277 KTAALKQVTAATFVTAHQIQEONANTLQAMHGLRQSTQDAATKANSKYINEDQPEQON 4336
QY 1317 HDA-----VANDEKVSIGE-----VNSLOSEM-----LGERDELOTSCALVS 1355
DB 4337 YDQAVQAAANNITIEQATLIDNNALINQAAATVNTTKAALHGDVKLQNDKDH-----AKQTVS 4392
QY 1356 EL-----ELLRAHVKEGEGNEITKGLNGLEKEIYIK--SESEVYK 1396
DB 4393 QLAHLNNAQKMEHTLIDSETTRTAVKQDLTEQAQLDQMLQALQSIDKQATRASSAYV 4452
QY 1397 SMLENLKEEDNKLKEQAEYSKEQNFLEE--VFSSQKQLVDETEVYKQALKAEEBLE 1454
DB 4453 NAEPRKQSYDEAYONASITAGLNPNITKNGVSSATQAYISSNALLDQVYERLAQK-- 4510
QY 1455 IKDRDYELVQATNTNLEKLETPPLQADHEEDSIDRSEMEIKVLGEKTERNOYLLER 1514
DB 4511 -----QTAGSNLHDLQLPRAQOQALENGQNNATNGEV--AQKLEBAQALMQA 4557
QY 1515 LOEKLELSKLEIIOKMEFVSYLKLDQLOK--LESLSBENIILKKNIDITLKHNSDT 1571
DB 4558 ME-----ALRNSIQDQOQTEAGSKFINEDKPOKDAVQAQVNAKDLINQTNPTL-----D 4608
QY 1572 QAOLOKTOQEQALKNLALASDNCPTIQEKEISADQVCHPEEKILLTEELHDKTQNEQ 1631
DB 4609 KAOVEQLTOAVNAQK-----DN--LHGDKLADDKQHAATD-----LNLQNGLNNQOR 4654
QY 1632 KLHLE-----KNELEQAQVELKC--EVEHLKMSIEMSKSLSLS--LOHEKHDEQEQ 1678

```

Db 4655 QALESOJNNATRGVNAQKLAFAKALDOAMQALNSIODQOQTESGSKFTNEDEKPOKDAY 4714
QY 1679 LILAKOQMOVTO-----EKKELQOJHEHLEAVDLKKNIELGMLFKNEAQOKTKEQ 1732
Db 4715 QAAVQNAKDLINOTGNPTLKSOVEQ---LTOAVTTAKNLEHODKLARDQOQAVTTVN 4770
QY 1733 CLINENKELSOHROCEIEELMKSLKDESALETTKESEOKVYNINQEMWLEMEE 1792
Db 4771 ALPLNNAQOQ-----ALTDALINAPTRTEVAOHV---QATLDEHMET 4812
QY 1793 LKNSQRTVIAERDQ-----LQDDRESVEMSIETODDLKRAQALQOQKVOELISQIS 1847
Db 4813 LKKNVDVNTDKAOPNTYEASTDKEAVDQALQAESITPTPTGNSANKADAVQV---LT 4869
QY 1848 VLEKISLE-NOMLYVANV-VKETLSERDDLN-----OSKOHU-----PSEIETIS 1892
Db 4870 KLOPEKELNGENEVAKTQAKOTIDQILHNLADQATKONIDQATKLOPIAEVDQA 4929
QY 1893 LSIKEFEPALEQAEKRAADARKTIDITERISNIE-----EOLLOQATNLKETLY 1942
Db 4930 TOLNOSMDLOQAVNEHAN-VEGTVDYTOADSKONAYKQAIADAENVLKNANKQOVDQ 4988
QY 1943 ERESLQCK-----EOLAL--NTEHLRETIKS-----KDLALGKMEQERD-----EAA 1993
Db 4989 ALQNIINAKQALNGDERVALAKTNGKHIDQILNNAQODGFKGRIDQSDNLQIQIIV 5048
QY 1984 NKVALTEKSSLEEQINENVTTLKEGEKEFPYLOPSSKOSSSOMEELRESLKTDL 2043
Db 5049 DEKALNRADQOLSQELTDN-----EGRTGSGTYV-----NADQVQKV-----YDE 5091
QY 2044 QLEBAKEISEBATEIKNLTK-ISSLEELIIONASTINEAVSERENLRSKOOLVSELE 2102
Db 5092 TVPKAKQALDKSTGO--NLTKAQVYKINDAV---TAAKALNGEERLNKKAALORLD 5145
QY 2103 QLS-LTKSRDHA------OSKREKDEAVNKIASLAEIKILKKEDEFR----- 2147
Db 5146 QLTHLNNAQOLATQOINNAETLKNASRAIRATKILDNAMGAVQYIDEQHLGYISSTNY 5205
QY 2148 -----DSKESLOEQSSHISEL-----CYKTELQMLKQOKEDINN-----KLAEKV 2190
Db 5206 INADDNLKAVNDANAIANAHELDKVOGNALAKAEQOLKONIIDQANALNGDOLANAKD 5265
QY 2191 EVDELLOHLSLKEQLODIQIOMELNEKRYTE-LCEKMDIMEKEI-----SVLRIMON 2242
Db 5266 KANAFVNSLNGLNOQ---OODLAHKAINNADVSDVTDIYNQIDLDAMETLKLHYDN 5321
QY 2243 E-POOEED-----DVAERM--DILESNOEIOELMEKISAVYSEQHTLSS 2285
Db 5322 EIPNAEQTVYQONADNAKTNFDKAKRLANTILNSDNTYNVDINGALQAVNDAIHN--- 5377
QY 2286 LSSELOKRETAHKHMLNIKESLSLTSRSFGLQTEHVKLNTOLTLNKFVYVYRTAA 2345
Db 5378 -----LNGQORLODARKAIQISIN-----QALANKLKEIEASNA 5411
QY 2346 VKEDHSIKYEDLA-----AEQKHDELRLQOLCLEQGRKWSASASEELKFCIEF 2399
Db 5412 TDQDKLAKKAEBELANSIINNINKATSNQAESQVQTAGNH-----AIEOVHANEI-- 5462
QY 2400 LNELLEFKKANIOSVODFSEVQVFLNOV-----GSTLOEELHKKGFOMLEEGDLHYD 2455
Db 5463 -----PKAK-IDANKYVDKQVALIDEIDRNPULTKEKQALKDRINOIILQO--GHNGIN 5514
QY 2456 AKKLSEGMQOENRIASTIOLTKRLKAVVQSK--TOREITVYLNQFEARKLOEKKEONKE 2513
Db 5515 NAMTKKEIEQAKQALQALQIDKIDLVAKEDAKQDVQVALIDEI-----DQNP 5566
QY 2514 LMRMEHNGSASVMEENARLLGILKTVODESKLOSRIKML-----ENELN--LYKDD- 2566
Db 5567 L-----TDKEQALKDRINOIILQOGHNDINNALTKEEI 5599
QY 2567 AMHGEKVALIIO--KILSRNAEELNAMOYKLTGKODNLOA-----AMKEIENL 2614
Db 5600 EQAKQALQALQIDKIDLVAKEDAKKAIKALANAKRQOINSNPULTPEQAKAKLKEIDEA 5659

```

```

QY 2615 OKMAYKAVPYKEEIDL-----KTKVYKLEMEKI--KYSKATDOEIAVLKSC 2660
Db 5660 EKRALDN-VENAOQTIDOLNGLNLGDDIDINTHWEEVDEQPAVNEIFENPBOIIVNGL 5718
QY 2661 LEDKEEGLR-----RLKEELRRAOADNTTVCVPRKYQAKASTPPVTCGGSGIVOST 2712
Db 5719 IYHRDITITBODILAHINLIDQLSAEVIDPSPYATISDSL--TAKVEVTLIDSSKYIVNV 5776
QY 2713 AMVLOSE-----KALERELSHYKKYHHLSTMSSEDRK-----KTRAKGDHS 2759
Db 5777 PVKVEKELSVVRQOALESTENAAQKINEINNSVTLTLEQKPAALAEVNLKLOQALIDHY 5836
QY 2760 SHTGSSHRGSPHKTETRYRH-GPYTPERSEMPSLHLSPKKSESSTRVYSPNRSIYSOL 2818
Db 5837 NDAPDVHSVEIQQOQBQAHIEQFNPEQFTIEQAKSNAKSIEDALQHM-----DEIKART 5892
QY 2819 VMSPGKTGMKHILSPSKVGLHKRRALSPNRSEMPPIOHYISPKGTGLHKMLTSTLFDNL 2878
Db 5893 DLTDKE-----KOEALIA-KLNOLEKEQAI-----QAIQRAQSIDEI 5926
QY 2879 SSPCKOQKVOENLNSPKGLFDVYSKSMPCPSOFPDNSKLGDF--SELNTAESNDKSOA 2936
Db 5927 SEOLEQPKAOKKAPLPAKEL-AKKQOEAISRKDSNEKINSIRSEIGTAD--EKQAA 5983
QY 2937 ENMWYEAKEET 2947
Db 5984 MNQINEIYLET 5994

```

Search completed: May 23, 2003, 15:03:20
 Job time : 340 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 23, 2003, 14:51:55 ; Search time 305 Seconds

(Without alignments)
2970.240 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769

Sequence: 1 MSEGDAVKVCVRVRLPIQRE.....QAENWVYAKKETAPBECKTS 2954

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-xip
-Q/cgn2_1/USFTO.spool/US09150867/runat_21052003_165913_20985/app_query.fasta.1.3143
-DB-issued.patents.NA -QPM-fastat -SUFFIX-rni -MINMATCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST-45 -DOCALIGN-200 -THR-SCORE-pct -THR-MAX-100 -THR-MIN-0 -ALIGN-15
-MODE-LOCAL -OUTFMT-pct -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USR-US09150867 @CGN.1.1.120 @runat.21052003.165913.20985 -NCPU-6 -ICPU-3
-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSYBLCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCPUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1153	7.8	10136	1 US-08-353-700-2	Sequence 2, Appl1
2	1153	7.8	10136	1 PCT-US95-16216-2	Sequence 2, Appl1
3	1061	7.2	8789	1 US-08-328-254-5	Sequence 5, Appl1
4	1061	7.1	4757	4 US-08-572-191-1	Sequence 1, Appl1
5	1051	7.1	4757	4 US-09-723-262-1	Sequence 1, Appl1
6	1051	7.1	4757	4 US-09-723-219-1	Sequence 7, Appl1
7	941	6.4	4127	4 US-09-592-054-7	Sequence 1, Appl1
8	915.5	6.2	4176	4 US-09-722-139-1	Sequence 1, Appl1
9	915.5	6.2	4176	4 US-09-721-832-1	Sequence 1, Appl1
10	915.5	6.2	4176	4 US-09-721-889-1	Sequence 1, Appl1
11	909.5	6.2	4308	4 US-09-592-054-1	Sequence 1, Appl1
12	867.5	5.9	3572	2 US-08-713-815A-2	Sequence 2, Appl1

13	866	5.9	30549	4 US-09-134-001C-322	Sequence 322, App
14	848	5.7	6773	4 US-09-166-350-27	Sequence 27, Appl
15	841.5	5.7	3837	4 US-09-724-517-1	Sequence 1, Appl1
16	841.5	5.7	3837	4 US-09-641-807A-1	Sequence 1, Appl1
17	841.5	5.7	3837	4 US-09-723-096-1	Sequence 1, Appl1
18	776	5.3	8948	4 US-09-643-597-119	Sequence 119, App
19	769.5	5.2	11091	4 US-09-134-001C-2243	Sequence 2243, Ap
20	759	5.1	3319	1 US-08-006-676B-2	Sequence 2, Appl
21	759	5.1	3319	1 US-08-289-845-1	Sequence 1, Appl1
22	759	5.1	3319	2 US-08-428-414A-4	Sequence 1, Appl1
23	759	5.1	3319	5 PCT-US94-00324-2	Sequence 4, Appl1
24	758.5	5.1	3709	4 US-09-541-782-7	Sequence 7, Appl1
25	758.5	5.1	3709	4 US-09-723-820-7	Sequence 7, Appl1
26	756.5	5.1	4911	4 US-09-718-692-1	Sequence 1, Appl1
27	756.5	5.1	4911	4 US-09-718-852-1	Sequence 1, Appl1
28	756.5	5.1	4911	4 US-09-718-815-1	Sequence 1, Appl1
29	754	5.1	13977	4 US-09-484-970B-60	Sequence 60, Appl
30	748.5	5.1	6306	1 US-08-195-487-3	Sequence 3, Appl1
31	748.5	5.1	6306	5 PCT-US93-06160-3	Sequence 3, Appl1
32	744.5	5.0	6306	1 US-08-466-390-3	Sequence 3, Appl1
33	744.5	5.0	6306	1 US-08-470-950-3	Sequence 3, Appl1
34	744.5	5.0	6306	1 US-08-467-781-3	Sequence 3, Appl1
35	744.5	5.0	6306	2 US-08-483-924-3	Sequence 3, Appl1
36	742	5.0	13307	3 US-08-836-022A-10	Sequence 10, Appl
37	742	5.0	13307	4 US-09-427-048A-10	Sequence 10, Appl
38	733	5.0	4223	4 US-09-541-782-5	Sequence 5, Appl1
39	733	5.0	4223	4 US-09-723-820-5	Sequence 5, Appl1
40	731	4.9	5093	1 US-08-468-036-23	Sequence 23, Appl
41	731	4.9	5093	2 US-08-376-843-23	Sequence 23, Appl
42	724	4.9	1421	4 US-09-592-054-5	Sequence 5, Appl1
43	724	4.9	1569	4 US-09-592-054-3	Sequence 3, Appl1
44	717.5	4.9	9551	1 US-08-056-200-93	Sequence 93, Appl
45	717.5	4.9	9551	2 US-08-800-644-93	Sequence 93, Appl

ALIGNMENTS

RESULT 1
US-08-353-700-2
; Sequence 2, Application US/08353700
; Patent No. 5599919
; GENERAL INFORMATION:
; APPLICANT: YEN, TIMOTHY J.
; APPLICANT: RATNER, JEROME B.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DANN, DOFFMAN, HERRELL AND SKILLMAN
; STREET: 1601 MARKET STREET, SUITE 720
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,700
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, JANET E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10136 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
US-08-353-700-2

Alignment Scores:
Pred. No.: 3,18e-83 Length: 10136
Score: 1153.00 Matches: 736
Percent Similarity: 37.84% Conservative: 581
Best Local Similarity: 21.15% Mismatches: 1163
Query Match: 7.81% Indels: 1001
Gaps: 144

US-09-150-867-1 (1-2954) x US-08-353-700-2 (1-10136)
QY 81 PheAlaTyrGlnThrSerSerGlyLysThrTyrThrMetGlyThrProAsnSer 100
Db 465 TTCACGAGAGACACGATTCAGCAAA----- 494
QY 101 LeuGlyTleIleProGlnAlaIleGlnGluValPheLysTleIleGlnGluIleProAsn 120
Db 495 -----AAACAATAGAAAAAAGTGGAGACGGAATTAAGG 530
QY 121 ArgGluPheLeuValSerTyrMetGluIleTyrAsnGluThrValLysAspLeu 140
Db 531 TGTAAATCTGAGCTTGAAGAACCAACAGCGCGAGCTGCAGATGCTCTGCAAT 590
QY 141 LeuGlyAspAspArgGlyLys-----LysProLeuGluIleArgGluAspPheAsn 157
Db 591 CCATGGAAATCACACCAAAAAATTTTACAACTCCACTACCAACCAATATATATAGT 650
QY 158 ArgAsnValTyrValAlaAspLeuThrGluGluLeuValMetValProGluHisValIle 177
Db 651 GGTTCACAGTAT---GAGATCTTAAGAAATAATATATAAGAGTTGAA----- 688
QY 178 GlnTrrPleLysGlyGluLysAsnArgHisTyrGlyLysGluThrLysMetAsnAspHis 197
Db 699 -----GACGAAAAAAGATAGAGCAGACAGAGTTAAAGCCTTGACAGCT 740
QY 198 SerSerArgSerHisThrIlePheArgMetIleValGluSerArgAsp-----ArgAsn 215
Db 741 AAAAAAGCAGACGACTCTTCACACAGCCACCATGAATCACCGCATTTGCCCGCAT 800
QY 216 AspProThrAsnSerGlu---AsnGlyAspGlyAlaValMetValSerHisLeuAsnLeu 234
Db 801 CACGCTTCATCATCTGTCTCTCATGCGACACAGAGAAGACCCCACTCATCTTTCA--- 857
QY 235 ValAspLeuAlaGlySerGluArgAlaSerGlnThrGlyAlaGluGlyValArgLeuLys 254
Db 858 -----TCATATCTCTCAAGAACT----- 875
QY 255 GluGlyCysAsnIleAsnArgSerLeuPheIleLeuGlnValIleLysLysLeuSer 274
Db 876 -----CCAATTAGAGAGATTTCTCT 896
QY 275 AspGlyGlnAlaGlyPheIleAsnTyrArgAspSerLysLeuThr----- 290
Db 897 GCATCTTACTTCTTGCGGAGACTAGAGTGACTCCAAAGTCAGATCACTTTGCCAATAGG 956
QY 291 ---ArgIleLeuGlnAsnSerLeuGlyLysAsnAlaLysThrValIleIleCysThrIle 309
Db 957 AAAAGAGATGCTAATAGCACTTCTTTGGCAATTTGAC----- 995
QY 310 ThrProValSerPheAspGluThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHis 329
Db 996 AGTCCTCATCTTTGGAT-----CAATTAAGCGCGAATCAAGAG 1037
QY 330 ValArgAsnThrProHisValAsnGluValLeuAspAspGluAlaLeuLysArgTyr 349

```

```

Db 1038 CTAGAAAC-----AAGATTAATGACTTG-----GAATACCGCTTCGCAAGACAT 1082
QY 350 ArgGlyGluIle-----LeuAspLeuLysGlnLeuGlu--- 361
Db 1083 GAAAAAGAAATGAAGGCCAAGTGAATAGTTTCAAGAACTCCAACTCCAACTGAGAA 1142
QY 362 ---AsnLeuGluSerSerGluThrLysAlaGlnAlaMetAlaLysGluGluHisThr 380
Db 1143 GCAAAAGTGAATTAATTTGAAAAAGAAAGTTTGAACAAATGTAGGAGTGAAGT 1202
QY 381 GlnLeuLeuAlaGluIleLysGlnLeu-----HisLysGluArgGluAspArgIle 397
Db 1203 AGAACACACACACAAATACACACCGGCTCACCAAGTATCTCATTTGAAACAAACAC 1262
QY 398 TrpHisLeuThrHisAsnIleValAlaLysSerGlnGluSerGln----- 413
Db 1263 AAAAAATGACGAGATTTGAGTTGTACAGCACAAATGACAAAGTGCAGATGTCT 1322
QY 414 ---AspGlnArgValLysArgValThrTrrAlaPro---GlyLysIleGln 431
Db 1323 CTGGAACAGAAAAATTAAGAAAAAGAGACTTTCACAGAGCTTCCCTCAACAG 1382
QY 432 AsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeu-ProGlyAs 451
Db 1383 CGTTCTTCCAAACACTGACCGACGAGTGCATCCAGTGAAGCCAGACTCACCCAGAG 1442
QY 451 nPheSerLysLysAlaLysPheSerAspMetProSerPheProGluIleAspAspSerGly 471
Db 1443 TTACAGCAAG-----CCAAGAAAT 1460
QY 471 LysThrGluPheSerAspPheAspAlaLeuSerMetMetAspSerAsnGlyIleAs 491
Db 1461 ATGCACACAGCTCTGACGCTGACGATGAATAACTACATCACTAAAGACGCTAGAA 1520
QY 491 PAlaGluTrrPAsnLeuAlaSerLysValThrHisArgGluLysThrSerLeuHisGlnSe 511
Db 1521 AACAAATTTGAAAGAGTTTAAGCAAAAGTGTGC-AGAGCTGAAGAGCGCTTCCAGCGAG 1579
QY 511 rMetIleAspPheGlyGlnIleSerAspSerValGlnPheHisAspSerSerLysGluAs 531
Db 1580 TCAGATCAAGAGAAATGACCTGAGAGAGCAAGAGAG-----GAATGAAAGAGAA 1633
QY 531 nGlnLeuGlnTyrLeuProLysAspSerGlyAspMetAlaGluCysArgLysAlaSerP 551
Db 1634 CAACCTCTTAAGAGTCACTCTGACAAAAGGCCACAGAAAGTGCACCTGCAGAGCAGA 1693
QY 551 eGluLysGluIleThrSerLeuGlnGlnLeuGlnSerLysGluGluLysLysG 571
Db 1694 ACTCAAGAACATCAACACAGTGTTAATCAGAGCCAGAAATTTGCAGAAAGAAATGAAG 1753
QY 571 uLeuValGlnSerPheGluLeuLysIleAlaGluLeuGluGluGlnLeuSerValLysAl 591
Db 1754 GAAGAAATACCTTCAGAGAAACCATGTTAAGAGATCTTCAAGAAAAATA----- 1802
QY 591 aLysAsnLeuGluMetValThrAsnSerArgGluHisSerIle----- 605
Db 1803 -----AATCAGCAAGAAACTCTTACTACTTGAAGAAACTGAA 1840
QY 606 -----AsnAlaGluValGlnThrAspValGluLysGluValAlaArgLysG 621
Db 1841 GCTTGCTGCTGCTGAGTGAAGAAAGCAGGAGATGTGTCACAGACCTTTGAAGAAAG 1900
QY 621 uMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSerSe 641
Db 1901 A-----GACATCCATTTGAACCACTTAATGATTAAGTAAAGACAGACAG 1945
QY 641 rValAspGlyLysArgLeuSerSerHisAspGluCysIleGlnHisArgLysMetIle 661
Db 1946 GAAGAGCTCCAAAGCCCTGTGAGTCT-----TT 1975
QY 661 uGlnGlnLysIleValAspLeuGluPheIleGluAsnLeuAsnLys----- 677

```

D	1976	AGAGTAAAAAAGAAACAAATATGCAAGATTGAAAGAAAGAAAGAAACTCTGTTTCTGTG	2035
O	678	-LysSerGluAsnAspLys-----GlnYSerSerGlnAspPheMetGluSe	694
D	2036	GAAGAAGTGAACCAACCACTTTTAACCTACAGATGCAATCAGAAAAGAAACTTCACAG	2095
O	694	r-----LleGlnLeuYs-----	698
D	2096	TAAATTTAATCAGCTGGGAAACTGTCTGAGACACAGCAAAATMAAAGTCATGATACAA	2155
O	699	-----GluAlaIleMetAlaGlnYsAlaSerAlaLeuGlnGluLeuAlaLeuMe	715
D	2156	CGAAGAGTAAAGAACCTGGAGAGTGAACAGACAGAAAACCTAATGTGCGAGATCAGAACT	2215
O	715	TArgAspAsnPheAspAsnIleIleLeuGlnAsnGlnIleuThrLeuYsArgGlnIleAlaAs	735
D	2216	TCACAACTGTGTAGACAGTAAGTCAAGTGAAGGAGAGACCCAGAAATCAGCTTATATGA	2275
O	735	PleuGlnArgSer-----	739
D	2276	GCTACACACAGAAAGCTGAGTTCACAGATCAGAAACATCAGAAAGAAATAGAAATATGTG	2335
O	740	-LeuYsGlnAsnGlnGlnIlePheGlnIleLeuGlnIleuGlnYsGlnIleThrGlnYsG	759
D	2336	TTTAAACATCTTCACAGCTACTCGGCAAGTGAAGATCTA-----GA	2377
O	759	uHISGluAlaGlnLeuIleHisGlnIleGlySerLeuYsLeuValGlnAsnAlaG	779
D	2378	ACACAACTTCAGTACTGTCTCA-----ATGAAATATATGACAAAGAACCG	2422
O	779	uMetYArgAsnGlnAsnLeuGlnIleuAspLeuGlnIleuThrYsThrYsLeuLeuYsGlnI	799
D	2423	GTCGTTAC--CAAGACTTGCATCCGCAAAATGAGACCTCAGAGGATCTGTAAATCCAA	2479
O	799	ngLIleGlnLeu-----	803
D	2480	AGATGCTTCTGTGTGACAAATGAATCATCAGAGAGCTTTTGCTTTGATCAGCA	2539
O	803	-----	803
D	2540	GCTGCATGCATCATCTCTTTCGCAAAATATATATGAGAAACAAGAACATGCCCTTACA	2599
O	804	----AlaGlnLeuArgLysArgAlaAspAsn-----LeuGlnY	815
D	2600	GAGGAGTGAATGTCGTTTAGAACACACCAAGATCCGAAAAATTCGCCATCTACACAAA	2659
O	815	sIysValArgAsnPheAspLeuSerValSerMet-----GlyAspSerG	830
D	2660	TAGAGTTGATTCACCTGAAATTTTCATTTAGAGTCTCAAAAACAGATGAACCTGACCTGA	2719
O	830	uLysLeuCYsGlnGlnIlePheGlnLeuYsGlnSerLeuSerAspAlaGlnAlaValTh	850
D	2720	AAACCATGTGGAAGAGTGTGTGCAAAATCAAGAGAGAAATAGAA-----GAAATCTCAT	2773
O	850	rArgAspAlaGlnYsGlnYsSerPheLeuArgSerGlnAsnLeuGlnLeuYsGlnY	870
D	2774	GAAAGCAGAACACAGTGCATCAAAAGTTTGTGGCTGAACAAAGTCAGCCATTAGTAGTT	2833
O	870	sMetGlnAspThrSerAsnTrpYsAsn-----G	880
D	2834	ACAGAGAAACACTTCTGCTCACAGAAATGTGTTGCTGAAACCTTAAGTCCCTTAGGAA	2893
O	880	nLysGlnLysAlaIaSerLeuPheGlnYsGlnLeuGlnIleuThrGlnYsSerAsnTrpY	900
D	2894	CAGAGAAAAAGAGCTGCACACTTTTAAATGATTAAGTAGAGAACTGACGCCAGAGATTCA	2933
O	900	sLysMetGln-----AlaAspLeuGlnYsGlnLeuGln-----	911
D	2994	AGAATTAATAAAGACAAACCATCTACTGTGAAGCTCTGTAAAGGAGCTACAACTTTATTC	3013
O	912	-----SerAl	913
D	3014	CGAAACCTTAGCTTGAGAGAAAGAAATATGATTCATCTTTAAATAAAGGGA	3073

QY	913	apheasncluleasnfyrtleuasnqlyleu	-----	alaglyly	927
					;
Db	3074	aatggaaagctggcccaagagatggactttaaagaaattaatgcattttaaataca	-----	glutengl	3133
QY	927	svajprokrgapleuenseratgval	-----	glutengl	939
Db	3134	agagaaagatgaacttaattccagaaanaagtgagaattttgcnaactatnagaatgaagga	-----		3133
QY	939	ulyslyvalserglubheserlysglneuglu	-----		950
					;
Db	3194	gaaagacatttcagagattactgtcagtcagacgaagaaaaactattttactcaaaag	-----	lyalaleuglgl	955
QY	951	-----	-----	lyalaleuglgl	955
Db	3254	atgcagaaacccggaattgcattgacgacttactcaaaaattacaaagccacacga	-----		3133
QY	955	ulysasnala	-----	leugluabngluvaltrhrcysleuserglutyrlyph	971
Db	3314	aaagaaattcttaatttaaatgactgtgtaaatgaaatccactactctttggaaataag	-----	3371	;
QY	971	eleuproasnluvalglucysleuysaanglilleserlysalaserglugluileme	-----		991
					;
Db	3372	-----	-----	-----	3427
QY	991	tleu	-----	leuysglnglucglunhisseralaserlleiserlysglgluilel	1010
					;
Db	3428	aaacaaatttaccattttcctgaagaaagaaatcagaattctgactgacagttgagacacgt	-----		3487
QY	1010	emecgl	-----	gluglaserglu	1016
					;
Db	3488	gcagacacctctgagatctgagatgacagataccaaacaaattcttaagacagcgctg	-----		3547
QY	1017	-----	-----	glutleuglneutrhraapgluvaltrhrlsthrnglserlyva	1032
					;
Db	3548	tgctttaagcaaaatctgactttaaagaaagaaacaaatgacaaagaaagact	-----		3607
QY	1032	1	-----	glnglthrnglucgluntyrtleuglmetlyslvsmethlaap	1047
					;
Db	3608	taatgacttaattacaaagagatgacacgctgaaagaaagtaagaaacttaaacatgaatg	-----		3667
QY	1048	-----	-----	aspleupheglustyrthleatgaaulysergluaglupleuleatrgl	1066
					;
Db	3668	tcaaaatttaaatcaaacagaaactttagaaactcttgaaagaaagaaagacagacagaa	-----		3727
QY	1066	umetgluasleuysgllythmetglu	-----		1075
					;
Db	3728	tcaatgtaattttaaaccctgagatgacttgaattaaagaaatttctgatatgta	-----		3787
QY	1076	-----	-----	seralgluvalyslilealaaptrhryslnglneuglgl	1090
					;
Db	3788	taatgcgactgttgcaatttaagacgtatgcataaataaagaaatttaaaccttcacga	-----		3847
QY	1090	utthrlleatgapyglnglneuleuhsiglyltyrphneghlnalamegl	-----		1110
					;
Db	3848	aagtgagaaagaaagagagcgctgcacagatga	-----	ttaca	3886
QY	1110	nthrlephroliethrproleuseraspsrleuproposerlysluvalglugl	-----		1130
					;
Db	3887	gacaaatt	-----	agagagagatcttgaamaaccacaaatttgcagagacat	3928
QY	1130	yasnserglasprolieglulileasnabytyr	-----		1141
					;
Db	3929	gcactgcacaaatattagtgcccttaaaagccttgaaatagatgcggagaaagactat	-----		3988
QY	1142	-----	-----	hisasnleuilealeualathrgluatgaaasnilemetvalcysle	1158
					;
Db	3989	ttcagggcctcatgattgttcacaaagctcaaaacgacaaatgcacaccttcagctctct	-----		4048
QY	1158	uglthrngluatgaaasberleuysglugluvalileaspleuasnthrglneuglnse	-----		1178
					;
Db	4049	gcaaaacaaatgaaacagctggaatg	-----	cttagagaaatattgtgaat	4096

QY 1178 rleuglnalaglnserlleugl-----lysersaspleugl 1190
 |||||
 Db 4097 ACTGACGCTGATAAAGATGAACTGTAAGTACAGCTGATGATTCAGAGTACAGATGAT 4156
 QY 1190 nlysrprolyslaspleuglnuglgluvallylsleuleuglnumetglule 1210
 Db 4157 CACAGCACTAGAAAATGGCAAGAGCTAGGCAAACTACTAATTAAGTAAATAT 4216
 QY 1210 ulysgl----- 1212
 Db 4217 AAATGATGACAGTGTCTCTCCATGGTGAAGTATAGTGAAGACATACCAGAGTGAATT 4276
 QY 1212 ----- 1212
 Db 4277 TGGTGAACAACCAATGATACAGACACCTGTGTCTTGGCTTCATTTGACAGAGTAATTC 4336
 QY 1213 -----HisleuthraspserserlnleuSerlleglulysleuglnleuglnleu 1230
 |||||
 Db 4337 CTACGACCACTTGACA-----TTGTGACAGCAAAAGAGTTCAAATGACACTTGGCGA 4387
 QY 1230 uvalthrglulysleuglnthrleuglnuglulmetlysasnlethrllleuglars 1250
 |||||
 Db 4388 ATTCAGAGAAATTTCTTACTTACAAAGTGAACAAATTTTACATGATCAGCACTG 4447
 QY 1250 ngululeuglnthrAsnpheluglspleuylsalaglunhisaspserserleuylsple 1270
 |||||
 Db 4448 TCACATAGCTCTTAAATGTCAGAGCTGACAGCTATGTTGACTCATTAAG----- 4499
 QY 1270 userglnasnllleuglnserlleglulthrlnasplulnleuarglnalagluln 1290
 |||||
 Db 4500 -GCCGAAATTTGGCTCTTGCA-----ACGAATCTGAGAACTTTCAAGGTGA 4546
 QY 1290 uleuarglnulnlyslleuvalaspserserphatrglnleuueuaspyserser 1310
 |||||
 Db 4547 CTGTGTGAAGAGATGACAGCTG-----GGCTTGAGAGAGGGCTGCTCCATCCCTGTC 4600
 QY 1310 lgllyleserserprosnhisasplavalalasnleuglnulysvalserleugl 1330
 |||||
 Db 4601 ATCTCTGTGTGCTGACAGCTCTAGCTTAC-----AGTTGGGAGA 4645
 QY 1330 uvalasnserleuglnserlleuuleuarglgluargaspulnleuglnthrsercy 1350
 |||||
 Db 4646 C-----TCCTCTTTTACAGACTCTTTTGA-----CAGACAGAGA 4684
 QY 1350 slysalaleuvalserleugln----- 1358
 |||||
 Db 4685 TATGTCTTTTGTAGTAATTTAGAGGGCTGTTTCAAGCAACAGTGCAGTATGATGA 4744
 QY 1359 -----LeuethrghlanhisvallysserValglulglunleugln 1374
 |||||
 Db 4745 AGTATTTTGCAGCAGTCTGCAGACCTATGTGTGACTCATTAAGGCCGAAATTTGGCTT 4804
 QY 1374 ethrlyslasleuasnlyleuglnulysgluleugllylserserlu-----gl 1391
 |||||
 Db 4805 GTCAACGAACTGTGAAATTTTCAAGGTGACTGTGCAAGAGATGAGAGTGGGCTTGA 4864
 QY 1391 userglnvalleuylsermetleuuglnleuylsasnlylsleuyls----- 1410
 |||||
 Db 4865 GGAAGGCTGTGTCATCTGTCTGTGTGCTGACAGCTCATGTCTTGAAG 4924
 QY 1411 -----gluglnalaglulthrser----- 1417
 |||||
 Db 4925 TTTGGAGACTCTCTTTTACAGACTCTTTTGAACACAGACAGATATGTCTCTTTT 4984
 QY 1418 -----Serlyslunsglnlnpheserleuuglnvalphe----- 1429
 |||||
 Db 4985 GAGTAATTTAGAGGGGTGTTTTCACCAAAACAGTGCAGTGTGATGAGATTTTTCAG 5044
 QY 1430 -----Serclyserserlnleuylsleuyls 1436
 |||||
 Db 5045 CAGCTGCAGAGAGAGATCTGACCAAGAAACCCCTTGGCCCCAGCGAGGAGGTGT 5104
 QY 1436 laspglulle-----gluvalleuylsalaglulnleuylsalagluln 1452

Db 5105 TGAAGACTGTAGTCCCTGTGTAGAGTGTACCCGACACTCCCTGAGAGACTAGAAAGAA 5164
 |||||
 QY 1452 gleugln-----lleysasparargpyrpheluleuvalglnthralas 1468
 |||||
 Db 5165 AATGGAAGTCAAGGATTTGAAAATTAAGAAATTCAGAACTGCAGCACTTATTAAG 5224
 QY 1468 nthrasnleuvalglu-----glulysleuglnthrproleuglnalaspri 1485
 |||||
 Db 5225 TTCTGAAGCAAGAGCTTGTACTGTGCAAGACAGTATTTGTGCAAAATATACAGCTG 5284
 QY 1486 -gluaspserserleasparargserleuglnumetglulileuvalleuylsle 1505
 |||||
 Db 5285 GCAACAGACACTGTGCAAGCTGACTGTGAGATGAGATGCCAAGTTGGCGCAGAAAAAGAA 5344
 QY 1505 ugluargasnlnthrleuugln-----Argleuglnleuglnleugl 1521
 |||||
 Db 5345 ACAGACGCAACACTGTCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5404
 QY 1521 uleuserasnlyleuglnulileuglnulysglumetglulthrserValleuyls 1541
 |||||
 Db 5405 CTTAAGTCTCG-----TCTTGTCTTGCAT 5431
 QY 1541 paspleuglnulnlyslleuglnserleuSerleuasnlelleuylsleuyls 1561
 |||||
 Db 5432 CGACACAGAGATGCTATTCAGAGC-----CGAATGAGAGCTG 5470
 QY 1561 easprlnthrleuylsnhisasparphrlnalaglulnleugln 1577
 |||||
 Db 5471 TGACATATCAAAAGATCTTCAAGAACTGACAGAAACCAACCAAGCATGATTTGA 5530
 QY 1578 -----Thrglnleuglnleuglnleu-----Alaly 1586
 |||||
 Db 5531 TCAGATTTGTGATTAAGATGCTGACAGAGCTCATTCATGACATGAGAAATTAAGTA 5590
 QY 1586 sasnlleuvallealalaserasparncysprolethrlnleuglnleu 1604
 |||||
 Db 5591 GACTGTGAGTGAAGAACCCAGAGAGTGTCTGGGAAAGTCCSSCATATTCAGATTA 5650
 QY 1605 -----Serlaasrcysvalnleuylsleu----- 1612
 |||||
 Db 5651 TGAGCTCCAGGGGAGATTAACCCAGAGCTCTTCAAGATGATTTGCAATTTGCAT 5710
 QY 1613 -----glulnlyslleu 1618
 |||||
 Db 5711 TTCTGTCTTAATGCTTGTGTGACTATGATTTCTGGGAAATCAGAGATATCCATGA 5770
 QY 1618 uleuthrlnleuglnleuylsleuylsleuylsleuylsleuylsleuyls 1638
 |||||
 Db 5771 TCTCAACTGTGGGTAAAGAGCATCAATGAGAAATTTGAGATTACTTCATGTGATGA 5830
 QY 1638 ngululeuglnalaglulnleuylsleuylsleuylsleuylsleuylsleuyls 1658
 |||||
 Db 5831 GAGCCTGACAG-----AAAGTTGAAGTCTTCTTAATGAAAT 5869
 QY 1658 ttleuglnserlysserserleuuglnleuylsleuylsleuylsleuylsleuyls 1678
 |||||
 Db 5870 GAAAGATTAAGTCAAACTC-----CATTTGACAGAGGTGA 5908
 QY 1678 nleuvala-----leuylsleuylsleuylsleuylsleuylsleuyls 1691
 |||||
 Db 5909 ACTAATGACCAAAATTTGAAGCATGCTAGATTTGGAATAATGATGGGAAATTAAGAA 5968
 QY 1691 ngululysgluleuglnlnthrlnleuylsleuylsleuylsleuylsleuylsleuyls 1711
 |||||
 Db 5969 AGAAACTCAGATTTAAGTGAATAATTTCTGTGATTCACAGAGATTAAG 6028
 QY 1711 sgluasnlleuglnleu-----glyleuasnphelulysasnleuylsleuyls 1728
 |||||
 Db 6029 CCAGAGAGTGAAGAACTTCTGAAGGCTCAATCTGATTTAGAAATGCATGACATTAATC 6088
 QY 1728 rthrlysglnleuylsleuylsleuylsleuylsleuylsleuylsleuyls 1739
 |||||

Db 6089 ATCAGCTGAAGATATGAGATTAATGTGCCAAGTGATGACAGCTGGAAGAGAGATT 6148
 QY 1740 -----GluLeuGlnInSerGlnHisArgLeuGlnCysGlu----- 1751
 Db 6149 TCTTGAATGGGAATAGAGCTGAGTAGATCAGATCGAGAAAGCTAGCATGAGCATGA 6208
 QY 1752 -----IleGlnGluLeuMetLysSerLeuLysAspLysGluSerLeuAlaLeuGluTh 1768
 Db 6209 AGCCCTTACTGAGAGCTGAGCTTAAAGAGTAGTCAAAACAGAGAAAGCATGTTTAAAGAA 6268
 QY 1768 rLeuLysGlnSerGlnGlnLysValIle---AsnLeuAsnGlnGlnMetGlnMetValMe 1787
 Db 6269 AGACAAATGAATTAAGCAAGAGTTATTGTCTGAGCAAGACACTCTCAGAGTGTAC 6328
 QY 1787 rLeuGlnMetGlnGlnLeuLysAsnSerGlnArgThrValIleAlaGluArg----- 1804
 Db 6329 AAGTGAGAGAAACAGCTTGTGTGAGAAATTAGATCTATGTCAAAAAAACACGAGCACT 6388
 QY 1805 -AspGlnLeuGlnAspAspLeuArgGlnSerValGlnMetSerIleGluThr----- 1821
 Db 6389 GGATCACTGTGTGAAAAAATGAAGAGAAACACAA---CAGCTGTGCTCATCAAG 6445
 QY 1822 -----GlnAspAspLeuArgLysAlaGlnGlnAlaLe 1832
 Db 6446 TGAAGTCTCCATTGCATTGAGTGCAGAGGAGGAGAGTGAAGGAAAAAGCGAACTCCT 6505
 QY 1832 uGlnGlnGlnLysAspLysValGlnGlnLeuThrSerGlnIleSerValLeuGlnGlnLys 1852
 Db 6506 TCAGACTTGTCTCTGTATGTAGTGTGAGTGTAAAGACAAACATATCTCCAGAGAA 6555
 QY 1852 sIleSerLeuLeuGlnLysAsnGlnMetLeuTyrAsnValAlaThrValLysGlnThrLeuSe 1872
 Db 6556 GCTGCAGAGAGTTGGAAAAAGCATCAGAGCATCTCTTGAACAAATGTAGCGTGGAAAA 6625
 QY 1872 rGluArgAspAspLeuAsnGlnSerLysGlnHisLeuPheSerGluIleGluThrLeuSe 1892
 Db 6626 CCAAAATTCGCAACTGAAATTAAGAGAAAGAAATTCCTGTCAAGAACTGAAACCTCGCA 6685
 QY 1892 rLeuSerLeuLysGlnLysGlnPhe-----AlaLeuGlnGlnAla 1905
 Db 6686 GGGCAGAGCTAGTGAATCAGATTTATGAAAAAGCTGATGTCTCCAGAGCTTGGAGCGGC 6745
 QY 1905 agLysLysAspLysAlaAspAlaAlaArgLysThrIleAspIleThrLysLysIleSerAs 1925
 Db 6746 ACGGTGTGAGAAAAGGTAGTTCGATGAGGCTGAGCTCAACAGAGAGAGAGCATCA 6805
 QY 1925 nIle-----GluGlnGlnLeuLe 1931
 Db 6806 GCTGAGAGAGGATCGAGAAACTGAGAGTTCGATGAGGCGCATGAAAAAGAGCAGCT 6865
 QY 1931 uGlnGlnAlaThrAsnLeuLysGlnThrLeuTyrGluArgGlnSerLeuIleGlnCysLys 1951
 Db 6866 GCACATCGCAGAGAACTGAAAGAACCGGAGGAGAAATGATTCACCT-----AA 6916
 QY 1951 sGlnGlnLeuAlaLeuAsnThrGlnHisLeuArgGlnThrLeuLysSerLysAspLeuAl 1971
 Db 6917 GGAATTAAGTGTGAGAACTTAAAGGAAATTGAGATGTCAAGAGAAACAGAGAGCTAGT 6976
 QY 1971 aLeuGlnLysMetGlnGlnLysArgGlnAlaAlaAsnLysValIleAlaLeuThrGln 1991
 Db 6977 GATT-----CTTGATGCCAGAAATTCACAAACA----- 7004
 QY 1991 uLysMetSerSerLeuGlnGlnGlnLysAsnGlnLysValIleThrThrLeuLysGlnGlnLys 2011
 Db 7005 -GAAGTAGAGACTCTAAAAACACAAATGAAGAGAGTCCAGAAAGCTGGAAGATTTTGA 7063
 QY 2011 u-----GlyGlnLysGlnThrPheTyrLeuGlnIleArgProSerLys 2024
 Db 7064 ATTAGACCTTGTACGTTAAGTGTGAAAAAGAAATCTACAAAAACAAATACAGAAAA 7123
 QY 2024 sGlnGlnSerSerSerGlnMetGlnLysArgGlnLysSerLeuLysThrLysAspLeuGln 2044
 Db 7124 ACAAGGTCACTGTGACAACTAGACAAAGTTACTCTTCAATTAAAGTCTGTAGAGAA 7183
 QY 2044 nLeuGlnGlnAlaGlnLysGlnLysSerGlnAlaThr----- 2056
 Db 7184 AAAGAGCAGACAGAGATGATGATCAAGAAATCTAAAGAACTCAGTGCAGATGCTTCA 7243
 QY 2057 -AsnGlnLysAsnLeuThrAlaLysIleSerSerLeu-----GluGlnGlnIleLe 2074
 Db 7244 GAATCAGTTAAAGAGCTAAATGAGGAGTACAGACCTTGTGTGTGTGCCAAAGAAATTA 7303
 QY 2074 uGln---AsnAlaSerIleLeuAsnGlnAlaValSerGlnArgGlnLysAsnLeuAlaGlnIleSe 2093
 Db 7304 GAAGGCCACAGACAGAGTCTTAGACCCACCAATAGAGAGAGAGATCAGCTGAGAAATAG 7363
 QY 2093 rLysGlnGlnLeuValSerGlnLeuGlu-----GlnLeuSerLeuThrLe 2108
 Db 7364 CATGMAAGCTGAGAGACCCCTAGAAAGCTGATGAAAAAGAAAGACCTCTGTCTTACA 7423
 QY 2108 uLysSerArgAspHisAlaPheAlaGlnSerLysArgGlnLysAspGlnAlaValAsnLys 2128
 Db 7424 ACAACTGAG-----GAAAGTGACCATCATGACATTTACTTAAGGTAG 7468
 QY 2128 sIleAlaSerLeuAlaGlnGlnLysIleLeuThrLysGlnMetAsp----- 2144
 Db 7469 AGTGAGAACCTTGAAGAGAGCTAGAGATGACAGAGCAAAACAGAGCATGACCTCT 7528
 QY 2145 -GluPheArgAspSerLys-----GluSerLeuGlnGlnInSerSerHisLeuSe 2161
 Db 7529 TGAGGAGAGAAATTCACAAAGAGAGTGAAGACCTTAACCAAAATAGAGGATGAC 7588
 QY 2161 rGlnGlnLeuCysThrTyrLysThrGlnLeuGlnMetLeuLysGlnGlnLysAspIle 2181
 Db 7589 CCAAAAGTGTGAGAGTCTGGAATTAAGATGTTGTACTATATAAGTACAGAAAAAGAAATCT 7648
 QY 2181 eAsnAsnLysLeuAlaGlnLysValLysGlnValAspGlnLeuLeuGlnHisLeuSerSe 2201
 Db 7649 GACAAATGAATTAACAAAGAGAGAGAGCCGAAATATCTGAAATA----- 7691
 QY 2201 rLeuLysGlnGlnLeuAspGlnIleGlnMetGlnLeuArgAsnGlnLysLeuArgAsnTyr 2221
 Db 7692 -----GAAATATTAATCATCATTTGAAAT-- 7718
 QY 2221 rGlnLeuCysGlnLysMetAspIleMetGlnLysGlnLysLeuArgLeuMetGln 2241
 Db 7719 -----ATTTCGCAAGAAAAAGA 7725
 QY 2241 nasGlnPheGlnGlnGlnLysAspAspValAlaGlnArgMetAspIleLeuGlnSerArg 2261
 Db 7736 GCAAGAGAAAGTACAGATGAAAGAAATCAAGACATGCTGAGAGATGCTTCAACAA- 7793
 QY 2261 gAsnGlnGlnIleGlnGlnLeuMetGlnLysIleSerAlaValLysSerGlnGlnIleThr 2281
 Db 7794 -----CAATTAAGAGCTCAATGAGAGAGTGCACCTGCAATATGCAACAAAGC 7846
 QY 2281 rLeuLeuSer-----SerLeuSerSerGlnLeuGlnLysGlnThrGlnAlaHisLys 2298
 Db 7847 CTGTAAAGGCCAAAGAGCAAGATCTTAGTACAGTAGAG----- 7886
 QY 2298 sHisCysMetLeuAsnIleLysGlnSerLeuSerSerThrLeuSerArgSerPheGlySe 2318
 Db 7887 ---TGCTT-----GAACCTTGAGAGAGCTCAGTTGCTTACAGAGCTTGTATGA 7930
 QY 2318 rLeuGlnThrGlnLysValLysLysAsnThrGlnGlnGlnThrThrLeuLysAsnLysPheLys 2338
 Db 7931 GGGCAAAATTAATTAATTTGTTTGCATATCTTCAAGTAAAGGCTCATTCAGAA- 7965
 QY 2338 sValValTyrArgThrAlaAlaValLysGlnAspHisSerLeuIleLysAspTyrGlnLys 2358
 Db 7966 -----CTAGAGAGTGCAGCA 8002
 QY 2358 sAspLeuAlaAlaGlnGlnLysArgHisAspGlnLeuArgLeuGlnLeuGlnCysLeuGln 2378
 Db 8003 GAAACTGTGAGAGAGAGATGAAGAAATCAGTAGACTGAAAAATCAAAATTCAGACCAAGA 8062

QY 2378 uGlnHisGlyArgLysTrpSerAspSerAlaSerGlnGluLeuLysPhe----- 2394
 DB 8063 GCAAGCTTGTCTCTAACTGTGCTCCAGTGGAGAGAGACCAACTTTGGAGAGAGCAAA 8122
 QY 2395 -CysGluIleGluPheLeuAsnGluLeuPheLysLysAlaAsnIleIleGlnSerVa 2414
 DB 8123 CTTAGAACACAGCAAACTGTGAGTGGAAATGGAGAGCAAGATCCAAATCTCAATCCAA 8182
 QY 2414 LglnAspAspPheSerGlnValGlnValPheLeuAsnGlnValGlySerThrLeuGlnG 2434
 DB 8183 AAATGCCCTTTTGCAGGACACATTAAGAGTGTGCAGAGTCTTCCAAAGAAATTAAGAAA 8242
 QY 2434 uGluLeuGlu-----HisLysGlyPheMetGlnTrpLeuGluGluPheGly----- 2450
 DB 8243 TGAGCTTGAATTCACAAATAATGACAAATAATGCTTTGTTGAAAAGTAACAAATAATGAC 8302
 QY 2451 -----AspLeuHisValAspAlaLysLys-----Le 2459
 DB 8303 TGCMAAGAAACTGACCTGCAGAGCAAAATGCATGAGATGCGACAGAAAACAGCAGAGCT 8362
 QY 2459 userGluGluMetGlnGlnGluAsnArgArgIleAlaSerThrIleGlnLeuLeuThrLy 2479
 DB 8363 GCACAGAGAACTACGTGAGAGAGCAAAATAGCTTACCTGAGAGATTCAGATTCTGTGGA 8422
 QY 2479 sArgLeuLysAlaValValGlnSerLysIleGlnArgGluIleThrValTyrLeuAsnG 2499
 DB 8423 AGAAATTAAGAGC-----AGCAAAAGATCAA----- 8447
 QY 2499 nPheGluAlaLysLeuGlnGluLysLysGlnGlnAsnLysGluLeuMetArgArgMetG 2519
 DB 8448 -----TTGAAAGAGCTCACACTAGAAATATGAAATTAAGAAAG----- 8486
 QY 2519 uHisHisGlyProSerAlaSerValMetGluGlnGluAsnAlaArgLeuLeuGlnIleLe 2539
 DB 8487 -----AGCCTAGATTCATGCACAAAGACCAAGAGCTGGAAGAGAGAG----- 8528
 QY 2539 uLysThrValGlnAspGluSerLysLysLeuGlnSerArgIleLysMetLeuGlnAsnG 2559
 DB 8529 -----AAAGTGAGAGAGAAATAGCTGAAATATCAGTACCG----- 8564
 QY 2559 uLeuAsnLeuValLysAspAspAlaMetHisLysGlyGluLys-----ValAlaIleLe 2577
 DB 8565 -----CTTCATGAGCTGAAAACCAAAACACACAGGCTTTGCT 8599
 QY 2577 uGlnAspLysLeuLeuSerArgAsnAlaGluAlaGluLeuAsnAlaMetGlnValLysLe 2597
 DB 8600 TTGGAC-----ACAACAAACAGTATGAAAGTAATCCACACATACGAGAGAAAT 8653
 QY 2597 uThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGluIleGlnAsnLeuGlnLysMe 2617
 DB 8654 GACTTCTTAAGAGAAATGTCTCAGTTCACAGAGCTGAGATAGACCTTTTAAAGTCTAG 8713
 QY 2617 tValAlaLysGlyAlaValProTyrLysGluGluIleAspAsn----- 2631
 DB 8714 T-----AAAGAGAGCTCAATTAATTCATTAAGAACTACTAC 8749
 QY 2632 -----LeuLysThrLysValValLysIleGluMetGluLysIleLysTyr----- 2646
 DB 8750 TCAGATTTTGGAGAAATGTAAGAAACCAAGATGACAAATCTAAATATGTAATCAGTT 8809
 QY 2647 -----SerLysAlaThrAspGlnGluIleAlaTyrLeuLysSerLys----- 2660
 DB 8810 GAAGAGAGAAATGAACGTCTCCAGGAGGAAATGAAGTTGTTCAATCAATCCGTGAACA 8869
 QY 2661 -LeuGluAspLysGluGlnGluLysArgArgLeuLysGluGluLeuArgArgIleGlnAl 2680
 DB 8870 GCTGGAAGAGGAAAGAGTACTGCAAGAAAGAACTCTCAACTTCACAGCTCAGACAGGA 8929
 QY 2680 aAspAsnAspThrThrValLysValProLysAspTyrGlnLysAlaSerThrPheProVa 2700
 DB 8930 GAAGCAGAAACAA----- 8942
 QY 2700 LThrCysGlyGlyLysSerGlyIleValGlnSerThrAlaMetLeuValLeuGlnSerG 2720

DB 8943 -----GTTACTGTATGATACCAAGGTCATGATTAACAACCTGA 8983
 QY 2720 u-----LysAlaAlaLeuGlu-----ArgGluLeuSerHisTyrLysLy 2733
 DB 8984 GATCAAGAACTGMAAGAAACTCTGMAAGAAAACCAAGAGGAGGAGATGAATACTTGA 9043
 QY 2733 sLysTyrHisLysLeuSerArgThrMetSerSerSerGluAspArgLysLysThrLysAl 2753
 DB 9044 TAGATCTGTCTCTTG-----CTATTAGCCCATGAAAAGTTAGAAAAGCTTAACA 9094
 QY 2753 aLysSerAspAlaHisSerSerHisThrGlySerSerHis-----ArgG 2768
 DB 9095 GATGTTAGAGACACAAAGTGGCCCATGTGTCTGCACAGCAATCAAAAGATTCGCCAGG 9154
 QY 2768 ySerProHisLysThrGlnThrTyrArgHisGlyProValThrProGluArgSerGlu 2788
 DB 9155 GTCTCT-----TTGCTAGTCCAGTTGTTCCAGGACCATCTCAAT 9196
 QY 2788 tProSerLeu-----HisLeuGlySerProLysLysSe 2799
 DB 9197 CCTTCTGTACTGAAAGAGTTATCATCTGSCCAAAATTAAGCTTCAGGCAAGAGGCA 9256
 QY 2799 rGluSerSer-----ThrLysArgValLysSerProAsnArgSerGluIleLy 2815
 DB 9257 AAGATCCAGTGAATATGGAGAAATGTGAGAGAGCACCAGACCTGATCCCAAGAGACT 9316
 QY 2815 rSer-----GlnLeuValMetSer-----ProGly-LysThrGlyMetH 2828
 DB 9317 TTTTAAAAAAGCAACAAAGACATCATGAGTGTATTCACCCCTGCAAGAAAGCA----- 9369
 QY 2828 sLysHisIleLeuSerProSerLysValGlyLeuHisLys-----LysArgAlaLeuS 2846
 DB 9370 --CGAAGAGTACTAGTTTGTAGCCAGAGGACTTCCAGAAATGTATAAAGAGGTGTGC 9427
 QY 2846 eProAsnArg-----SerGluMetProHisGlnHisV 2857
 DB 9428 TGACATCCCGACAGAAAGACTAGCCCATATATCTGCGAAGAACCAACATGGCA----- 9483
 QY 2857 aLleSerProGlyLysThrGlyLeuHisLysAsn-----LeuThrG 2871
 DB 9484 --CTCGACACAGCCCGCCGCTGCTGCACAGAAATTAAGGCTATCCCACTAGTGTCCG 9541
 QY 2871 LuSerThrLeuPheAspAsnLeuSerSerProCysLysGlnGlnLysValGlnGluAsn 2891
 DB 9542 CAAGAAGAAATC-----TTGCAGAGTCTCCAAACCAACAGCGTGTGCA-----G 9586
 QY 2891 euAsnSerProLysGlyLysLeuPheAsp-----ValLysSerL 2904
 DB 9587 CAGATCACAAAAGGTCAAGTTGCTCAGCGGAGCCCAAGTAGATTCAGGCAACATCTCCG 9646
 QY 2904 ySerMetProTyrCysProSerGln-----PhePheAspAsnSerLysLeuGlyA 2921
 DB 9647 AGAAGCCACACAGCAAAATCCGTCCTCAGTCAATATCTCTGAGAGAAAGTCCGACTGACAG 9706
 QY 2921 sPheSerGluLeuAsnThrAlaGluSerAsnAspLysSerGlnAlaGluAsnTrp 2939
 DB 9707 CCCCAAGAGAGGCGCTGAGGCTCAAGCAGAGCGCAGCTTGTCCAGCCCAAGGCTGG 9762
 RESULT 2
 PCF-US95-16216-2
 ; Sequence 2, Application PC/TUS9516216
 ; GENERAL INFORMATION:
 ; APPLICANT: Yen, Timothy J.
 ; APPLICANT: Ratner, Jerome B.
 ; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently
 ; TITLE OF INVENTION: Expressed Kinetochores Protein, and Methods of Use
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
 ; STREET: 1601 Market Street Suite 720
 ; CITY: Philadelphia
 ; STATE: PA

```

COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16216
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/353,700
FILING DATE: 09-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10136 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
PCT-US95-16216-2

```

Alignment Scores:

```

Pred. No.: 3,18e-83 Length: 10136
Score: 1153.00 Matches: 736
Percent Similarity: 37.84% Conservative: 581
Best Local Similarity: 21.15% Mismatches: 1163
Query Match: 7.81% Indels: 1001
DB: 5 Gaps: 144

```

US-09-150-867-1 (1-2954) x PCT-US95-16216-2 (1-10136)

```

QY 81 PheAlaTyrGlnThrSerSerGlyLysThrThrMetMetGlyThrProAsnSer 100
DB 465 TTCACAGAGGACACGATTCAGGCAAA----- 494
QY 101 LeuGlyIleLeuProGlnAlaIleGlnGluValPheLysIleIleGlnGluIleProAsn 120
DB 495 -----AAACAAATAGAAAACTGGACACGAACTTAAAGG 530
QY 121 ArgGluPheLeuAlaArgValSerTyrMetGluIleTyrAsnGluThrValLysAspLeu 140
DB 531 TGTAAATCTGAGCTTGAAGAGACCAAGCGCGAGCTCGACAGTGTCTCTCAAT 590
QY 141 LeuCysAspAspArgArgLys-----LysProLeuGluIleArgGluAspPheAsn 157
DB 591 CCATGGAAATCACCAAAAAATTTTACAACTCCACTTACACCAAGTCAATATTATAGT 650
QY 158 ArgAsnValTyrValAlaAspLeuThrGlnGluLeuValMetValProGluHisValIle 177
DB 651 GGTTCAGAGTAT--GAGAGTCTAAAGAAAAATATATATAAAGAGTTGAA----- 698
QY 178 GlnTrpIleLysGlyGluLysAsnArgHisTyrGlyGluThrLysMetAsnAspHis 197
DB 699 -----GACGAAAAAAGATTGAGGACAGAGGTTAAAGCTTCGACGCT 740
QY 198 SerSerArgSerHisThrIlePheArgMetIleValGlnSerArgAsp-----ArgAsn 215
DB 741 AAAAAAGCAAGCAAGCTCTTCCACCAAGCCACCATGAAATCAGCGCATTTGCCCGCAT 800
QY 216 AspProThrAsnSerGlu---AsnGlyAspGlyAlaValMetValSerHisLeuAsnLeu 234
DB 801 CAGCTTCATCATCTGTTCTCATGCGACAGAGAAAGACCCCAAGTCAATCTTTCA--- 857

```

```

QY 235 ValAspLeuAlaGlySerGluArgAlaSerGlnThrGlyAlaGlnGlyValArgLeuLys 254
DB 858 -----TCTAATTTCTCAAGAACT----- 875
QY 255 GlnGlyCysAsnIleAsnArgSerLeuPheIleLeuGlnGluValIleLysLysLeuSer 274
DB 876 -----CCAAATTAGAGAGATTTCTCT 896
QY 275 AspGlyGlnAlaGlyPheIleAsnTyrArgAspSerLysLeuThr----- 290
DB 897 GCATCTTACTTTTCTGGGAACTAGAGTGACTCCAAAGTCAATGATCAATTTGCAAAATAGG 956
QY 291 ---ArgIleLeuGlnAsnSerLeuGlyLysAsnAlaLysThrValIleIleCysThrIle 309
DB 957 AAAAGAGATGTAATAGCAGTTTCTTGGCAATTTGAC----- 995
QY 310 ThrProValSerPheAspGluThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHis 329
DB 996 AGTCCTCATCTTTTGAT-----CAATTAAAGCGCAGAAATCAAGAG 1037
QY 330 ValArgAsnThrProHisValAsnGluValLeuAspAspGlyAlaLeuLysArgTyr 349
DB 1038 CTAAGAAAC-----AGATTAAAGAGTTG-----GAACTACGCTGCAAGGACAT 1082
QY 350 ArgLysGluIle-----LeuAspLeuLysGlnLeuGlu--- 361
DB 1083 GAAAAAGAAATGAAAGGCCCAAGTGAATAGTTTCAAGACTTCCAACTCCAGTGGAGAAA 1142
QY 362 ---AsnLeuGlnSerSerSerGlyThrLysAlaGlnAlaMetAlaLysGlnLysThr 380
DB 1143 GCAAAAGTGAATTAATTAAGAAAAAGAAAGTTTGAACAAATGATGAGATGAACTAGTG 1202
QY 381 GlnLeuLeuAlaGlnIleLysGlnLeu-----HisLysGlnAlaGlnLysArgGly 397
DB 1203 AGAACAACAGACATATAGCAGCAGCGCTCAACCAAGATATGATGATTTGACAAACAAACTG 1262
QY 398 TrpHisLeuThrAsnIleValAlaLaserSerGlnGlnSerGlnGln----- 413
DB 1263 AAAAAATTTGACGAAAGATTGAGTTGTCAGCGACAAATGACGAAAGTCCAGATGTTCT 1322
QY 414 ---AspGlnArgValLysArgLysArgValThrTrpAlaPro---GlyLysIleGln 431
DB 1323 CTGACACAGAAATTAAGAAAAAGAAAGAGTTTCAAGAGGAGCTTCCCGTCAACAG 1382
QY 432 AsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeu---ProGlyAs 451
DB 1383 CGTTCTTCCAAACACATGACGACGACGAGTGCATCCAGATGAAGGCCAGAGCTCACCAGAG 1442
QY 451 nPheSerLysLysAlaLysPheSerAspMetProSerPheProGluIleAspAspSerVa 471
DB 1443 TTACGCAAG----- 1460
QY 471 LysThrGlnPheSerAspPheAspAlaLeuSerMetLysSerAsnGlyIleAs 491
DB 1461 ATGCACACAGCTCTGACGAGTGAATGAATCAATCATCATGTAAGCAACACCTGAA 1520
QY 491 PAlaGluTrpAsnLeuAlaSerLysValThrHisArgGlnLysThrSerLeuHisGlnSe 511
DB 1521 AACAAATTGGAAGAGTTTAAACAAAGTTGTC-AGAGCTGAACAGCGCTTCCAGCGGAG 1579
QY 511 rMetIleAspPheGlyGlnIleSerAspSerValGlnPheHisAspSerSerLysGluAs 531
DB 1580 TCAATGACAGAGAAATAGCTGAGAGAACATGAG-----CAATTAAGAAAGAGAAA 1633
QY 531 nGlnLeuGlnTyrLeuProLysAspSerGlyAspMetAlaGlnCysArgLysAlaSerPh 551
DB 1634 CAACCTCTTAAGAGTCACTCTGAGCAAAAGGCCAGAGAAAGTGGCCACCTGAGGAGAGA 1693
QY 551 eGluLysGlnIleThrSerLeuGlnGlnIleGlnLysGlnGluGlnGlyLysGln 571
DB 1694 ACTCAGAAACATCAACACAGTGTTAATTCAGAGCCAAATTTTGCAGAAATATGAAGAC 1753
QY 571 uLeuValGlnSerPheGluLeuLysIleAlaGlnGlnGlnIleuSerValLysAl 591

```

```

Db 1754 GAAGAAATACCTCTCAGGAAACCATTGTTAGAGATCTTCAAGAAAAATA----- 1802
QY 591 aLysasnLeuGluMetValThrAsnSerArgLuhISerLe----- 605
Db 1803 -----AATCACAAGAAAGAACTCTTGACTTTAGAAAAACTGAA 1840
QY 606 -----AsnAlaGluValGlnThrAspValGluLysGluValArgLysG 621
Db 1841 GCTTCGTGCTGCTGATCTGGAAGAACGAGATGTTCTCAAGACCTTTTGAGAAAAAG 1900
QY 621 uMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSer 641
Db 1901 A-----GAACATTCACATTGAACAACCTTAATATGATTAATGAACCAAGACA 1945
QY 641 tValAspGlyLysArgLeuSerSerSerHisAspGluLysIleGluLuhISargLysMet 661
Db 1946 GAAAGAGTCACAAAGCCTTGCTGAGTGT-----TT 1975
QY 661 uGluGlnLysIleValAspLeuGluGlnPheIleGluLuhISnLys----- 677
Db 1976 AGAGTTAAAAAGAAAGAAATATGAAGATTGAAGAAAGAAACCTGTTCTTCTGTTG 2035
QY 678 -LysSerGluAsnAspLys-----GlnLysSerSerGluGlnAspPheMetGlu 694
Db 2036 GAAAAAGTGAAGAAACCTTTAACTCAGATGATCAGAAAGAAACCTTGACAG 2095
QY 694 r-----IleGlnLeuLys----- 698
Db 2096 TAAATTAATTCACCTTGGAACCTGCTGGAAGACACAGCAATAAAAGTCATGATACAA 2155
QY 699 -----GluAlaIleMetAlaGluLysAlaAsnAlaLeuGluLuhISnLys 715
Db 2156 CGAGAGAGTAAAGAACCTGAGATGACAGACAGAAACCTTAAGTGTGAGATGAGAAACCT 2215
QY 715 tArgAspAsnPheAspAsnIleIleLeuGluAsnGlnLysLeuLysArgGluIleAlaAs 735
Db 2216 TCACAAAGCTGTGACACAGTAAGTCAGTGAAGAGAGAGAACCAAACTAGCTAATATGA 2275
QY 735 pLeuGluArgSer----- 739
Db 2276 GCTACACACAAAGACCTGAGTTCACAGATCAGAAACATCAGAACGAAATAGAAATATGTC 2335
QY 740 -LeuLysGluAsnGlnGlnLysThrAsnGlnPheGlnIleLeuGluLysGluLysGln 759
Db 2336 TTGGAAGACTTCTCAGCTTCTGCGCAAGTGAAGATCTA-----GA 2377
QY 759 uHISGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsnAlaG 779
Db 2378 ACACAAAGCTTCAGTTACTGTCA-----AATGAATAATGACAAAGACCG 2422
QY 779 uMetTyrAsnGlnAsnLeuGluGlnAspLeuGluThrLysThrLysLeuLeuLysGlu 799
Db 2423 GTGTATC-----CAAGACTTCGATGCCCAATATGAGAGCTTCAGAGATCTGTAAATCCAA 2479
QY 799 nGluIleGlnLeu----- 803
Db 2480 AGATGCTCTCTGTGCAAAATGAGATCATCAGAGAAGTCTTTGGCTTTGATCAGCA 2539
QY 803 ----- 803
Db 2540 GCCTGCATGCATCATCTCTTGCAAAATTAATTGAGAAACAAGAGACATGCCCTTCA 2599
QY 804 -----AlaGluLeuArgLysArgAlaAspAsn-----LeuGlnLys 815
Db 2600 GAGAGTGAATGCTGTTAGAGACAGACAAAGTCCGAAAAATTTGCCATCTTCAAAA 2659
QY 815 sLysValAlaArgAsnPheAspLeuSerValSerMet-----GlyAspSerG 830
Db 2660 TAGAGTGAATGCTGTTAGATTTTCAATAGAGTCTCAAAAAGAGATGACATGACCTCA 2719
QY 830 uLysLeuGlyGlnGluIlePheGlnLeuLysGlnSerLeuSerAspAlaGluAlaValTh 850
:||||| |||||||: |||: |||

```

```

Db 2720 AAAGCACTGTGAAGAGTGTGCTGCAAAATCAAGAGAAATAGAA-----GAAATCTCAT 2773
QY 850 rArgAspAlaGlnLysGluLysSerPheLeuArgSerGluAsnLeuGluLysGluLys 870
Db 2774 GAAAGCAGACAGATGATCATCAAAAGTTTGTGCTGAACAAGTCAAGCATTAAGTAAGT 2833
QY 870 sMetGluAspThrSerAsnTyrAsn----- 880
Db 2834 ACAGGAAGACACTTCTGCTCACCAAGATGTTGCTGAACCTTAAGTGCCTTGAGCA 2893
QY 880 nLysGluLysAlaAlaSerLeuPheGluLysGlnLeuGlnLysGlnLysSerAsnTyr 900
Db 2894 CAAGAAAAAGACCTGCAACTTTAATGATAGATAGAACTGACAGCAGACATTC 2953
QY 900 sLysMetGlu-----AlaAspLeuGlnLysGluLeuGln----- 911
Db 2954 AGATTAATAAAAGACCAACCATCTACTGTAAGACTCTCTTAAGAGACTTAATTC 3013
QY 912 -----SerAl 913
Db 3014 CGAAACCTTAAGCTTGAGAGAAAGAAATGACATTCATCTTTAATTAAGAGCA 3073
QY 913 aPheAsnGlnIleAsnTyrLeuAsnGlyLeu-----AlaGlyLys 927
Db 3074 AATGGAAGAGCTGACCCCAAGAGATGGGACTCTTAAGCAATTAATGCATCCTTAAT 3133
QY 927 sValProArgAspLeuLeuSerArgVal-----GluLeuG 939
Db 3134 AGAGAAAGATGAACCTTAATTCAGAAAGAGTGAAGCTTTGCCAACTATATATAGATGA 3193
QY 939 uLysLysValSerGluPheSerLysGlnLeuGlu----- 950
Db 3194 GAAAGACATTTCAAGATTAATCTGATCAGTACAGACAGAAAGAACTATTTACTACAA 3253
QY 951 -----LysAlaLeuGluG 955
Db 3254 ATGTAGAGAAACCGAAATGCATATAGAGATCTTAGTCMAAAATACAAAGCAGCA 3313
QY 955 uLysAsnAla-----LeuGluAsnGlnValThrCysLeuSerGluTyrLysPh 971
Db 3314 AAAGAAATCTTAATTAAGATGCTTCTTAATGAATGCATAGTCTTTGTGAATAATGAG 3371
QY 971 eLeuProAsnGlnValGluCysLeuLysAsnGlnIleSerLysAlaSerGluLysIle 991
Db 3372 -----AAAAATGAGTTGGAACAGCTAAAGAGAGACTTTGCCAAAGACCAAGAA 3427
QY 991 tLeu-----LeuLysGlnGlnGluGlnHisSerAlaSerIleLeuSerLysGlnGlu 1010
Db 3428 AACAAATTTAGCATTTGCTGAAGAAAGAAATCAGAAATCTGATGAGTGAAGACAGT 3487
QY 1010 eMetGln-----GluLysnSerGlu----- 1016
Db 3488 GCACAAAGCTGTGAGATCTGAGATGACAGATTAACCAAAACATTTCTAAGACGAGCT 3547
QY 1017 -----GlnIleLeuGlnLeuThrAspGlnValThrHisThrGlnSerLys 1032
Db 3548 TGGTTTAAGCAGAAATCTGACTTTAAAGAGAAACAAACAAATGCAAAAGAGAAAGT 3607
QY 1032 l-----GlnIleThrGlnGlnGlnIleTyrLeuGlnIleMetLysLysMetHis 1047
Db 3608 TAATGACTTATTAACAGAGATGACAGACGTCGATGAAGAGTAATGAAGACTTAACAT 3667
QY 1048 -----AspLeuPheGlnLysTyrIleArgAsnLysSerGluAlaGlnAspLeuArg 1066
Db 3668 TCAAAATCTGAATTCAGAACCAATTTAGAACTGTGAAGAAAGAGAGAGTGAAGAAA 3727
QY 1066 uMetGluAsnLeuLysGlyThrMetGlu----- 1075
Db 3728 TCAATGAATTTTAACCTGATGATGATCTTGAAGTAAGAAATTTCTGTAGATAGT 3787
QY 1076 -----SerValGluValLysIleAlaAspThrLysLysGlnLysGlnLys 1090
Db 3788 TAATGCGCACTGTGCTCAATTAAGACSTATGCTAAGAAATTAAGAAATTAACCTGAG 3847
:||||| |||||||: |||: |||

```


OY	1090	uThrIleArgAspIysGluIleuLeuThsIsgLulysIySTyrrPhepneGlnAlaMetGI	1110
Db	3848	AAGTGAAGAGAGAGAGAGAGCTGCAGCATGAA-----TTACA	3866
OY	1110	nThrIlePheProIleThrProLeuSerAspSerLeuProProSerLysLeuValGluGI	1130
Db	3887	GACAAAT-----AGAGAGACTCTTGAACACACACATTTTGCAAGACAT	3928
OY	1130	YAsnSerGlnAspProIleGluIleAsnAspIyr-----	1141
Db	3929	GCACCTACCAAGAAATTAAGTGGCTTTAAAGACCTGTGAATAGATGGGAGAAAGTAAAT	3988
OY	1142	-----HisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLe	1158
Db	3989	TTTCAGGCGCTCATGATGATTGTTCACCAAGTCACAAACGCAATGCAACCTTCAGTCTCT	4048
OY	1158	uGluThrGluArgAsnSerLeuLysGluGlnValIleAspLeuAsnThrGlnLeuGlnSe	1178
Db	4049	GCAACCAACAAATGACCAACCTGATGAG-----CTAGAGAAATAATGTGAAT	4096
OY	1178	rLeuGlnAlaGlnSerIleGlu-----LysSerAspLeuGI	1190
Db	4097	ACTCCAGCTGCAAAAGTATGATGCTGTAACTGATGCTGAATGATTCACGTGCAATGTAT	4156
OY	1190	nLysProLysGlnAspLeuGluGluGlyGlnValIlyLeuLeuLeuGluMetGluLeuLe	1210
Db	4157	CACAGCAACATGAGAAATGGCACAABAAGTAGAGAAACTCAATTAATGAAGTTAAATTT	4216
OY	1210	uLysGly-----	1212
Db	4217	AAATGATGACAGTGGCTTCTTCATGCTGATGATTAAGTGAAGACATACCAGAGGTGAATT	4276
OY	1212	-----	1212
Db	4277	TGGGAAACAACAATGAACAGACCCCTGTCTTGGCTCCATTGGACGAGAGTAATTC	4336
OY	1213	-----HisLeuThrAspSerGlnLeuSerIleGluLysLeuGlnLeuGluAsnLeuGI	1230
Db	4337	CTACGACACACTTGACA-----TTGTGACAGCAAAAGATTCAAAGCACTTTGGCCGA	4387
OY	1230	uValThrGluLysLeuGlnIleThrLeuGlnGluMetLysAsnIleThrIleGluArgAs	1250
Db	4388	ATTCAGAGAGAAATTTCTTATCTTTACAAAGTGACACCAAAATTTTACATGATCGACACTG	4447
OY	1250	nGluLeuGlnIleThrAsnPheGluAspLeuLysAlaGlnHisAspSerLeuLysGlnAspLe	1270
Db	4448	TCAGATGAGCTCAAAATGTCAAGACTGCAGACCTATGTTGCATTAATAAG-----	4499
OY	1270	uSerGluAsnIleGluGlnSerIleGluThrGlnAspGluLeuArgAlaAlaGlnGluGI	1290
Db	4500	-GCCGAAAAATTTGGTCTGTCA-----ACGAAATCTGCAAAACTTTCAAAGTGA	4546
OY	1290	uLeuArgGlnGluLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSerVa	1310
Db	4547	CTTGGTAAAGAGATGCAGCTG-----GGCTTGAGAGAGGGGCTCTGTCATCCCTGTC	4600
OY	1310	IleGlyLeuSerSerProAsnHisAspAlaValAlaAsnGlnLulysValSerLeuGlyGI	1330
Db	4601	ATCTCTTGTGTCTGACAGCTCTGATGCTTTAGC-----AGTTGGGAGA	4645
OY	1330	uValAsnSerLeuGlnSerGluMetLeuArgGlyGluArgAspGluLeuGlnIleThrSecY	1350
Db	4646	C-----TCTCTCTTTTACAGAGCTTTTAAAG-----CAGACGAGAGA	4664
OY	1350	sLysAlaLeuValSerGluLeuLeu-----	1358
Db	4685	TATGTCTCTTTGATTAATTAGAAGGGCTGTTTCAGCAAAACGAGTCAGATGATGA	4744
OY	1359	-----LeuLeuArgAlaHisValLysSerValGluGlyGluAsnLeuGluI	1374
Db	4745	AGATATTTTGGACAGAGCTGTGCAGACCTATGTGACATTTAAAGGGCCAAATTTGGCTCT	4804

QY	1374	ethrlnlylsylnuansnclyleuulnlysgulileuengllyrssegrlu-----gl	1391
Db	4805	gtccacagaaactctggaaaacttttaaacgtcttggtgaagagatgcacgtggccttggaa	4864
QY	1391	userglvalleuysersertleuqluansleuulysgluabapnaabnlylsleuys--	1410
Db	4865	ggagggcctcgtccatccctctgcacatcccttctgtgcctgcacacactcttactttacgc	4924
QY	1411	-----gluinalagluurysr-----	1417
Db	4925	tttggagactcctcttttacagacactctttttagacacagacagagatattgctctttt	4984
QY	1418	-----serlysgluansnglnpheserleugluqlualphe-----	1429
Db	4985	gagtaatttagaagggtgttttcaccaaaacagtcacatgttagatgaagatttttgcag	5044
QY	1430	-----serglysergluylsleua	1436
Db	5045	cacgtcgcagagagagaattctgacacagaaagaaaccccttgcggccacaggaagggtgt	5104
QY	1436	laspqluile-----gluvalleuysalaglnleuulysgluabalaagluuar	1452
Db	5105	tgaagacttgcagtcctctgtgaggtgacccgcagtcctctgcagaaagctagaagaa	5164
QY	1452	gleglu-----llelysapargaptryphleglualevalglnthralas	1468
Db	5165	aatggaaagtcamaagatttttggaaaaataagaaattcamaagctgcagcagttattttag	5224
QY	1468	nthrasluvalglu---glylsyleuulntrhproleuaglnalabaphlsglu-----	1485
Db	5225	tttcgaaggcagaaagcttactgcctttgagaagcagtttttgcagaaatgaacacgtg	5284
QY	1486	gluasperlleaapargarsergluwetgluileuyluileuylgluylsle	1505
Db	5285	gcacacagagctgcacagcctgactctggagatggagtcacagttggcggcagaaagaa	5344
QY	1505	ugluaragsnclntryleuenglul-----argleuenglugluylsleuyl	1521
Db	5345	acacagcgaacacactgtcactttgagctggagagacagacacccagctacacaggtcgga	5404
QY	1521	uleuserasnlylsleuulileuenglulysgluwetgluthrserValleuuleuylsas	1541
Db	5405	cttagttctcgcg-----tctttgctttggcat	5431
QY	1541	paspberlnclnlylsleuulserleuenseuergluanslleleuylsgluasnl	1561
Db	5432	gcacacagaaagatcctattccaaagc-----cgaaatgagagctg	5470
QY	1561	easpthrThrleuylshishiseraspthrclnalaaglnleuyls-----	1577
Db	5471	tgacatattcaaaaagacattcttcacaaaactacagaaacacccaaagcatgtattcca	5530
QY	1578	-----thrclnlnluuengluleu-----Alaly	1586
Db	5531	tcgaattttgcatTAAGATGCTAGAGAGACCTCAATCTAGACATTGAGAAATTAATCGA	5590
QY	1586	sasnleuAlailelaalelaseraspansprrlthrnglnluylsgluthr-----	1604
Db	5591	gacgtgtgcacagtgaaacccacagagaaagtcctctgggaaacagtcgccacagataccatTA	5650
QY	1605	-----Serlaaspcysvalhlsproleu-----	1612
Db	5651	tgaccttcacagggagagatTAACCCAGGGGCTTCAGAAATGCTATTGTAATTGTCTATT	5710
QY	1613	-----glugluylsThleule	1618
Db	5711	tttcgtgtcTAATGCTTTGGTACCTATGGAATTTCTGGGGAATCGAGAAATATTCCTAA	5770
QY	1618	uleuThrglugluuentslgnlyshrasnnglnuglnlylsleuuhlsiglulysas	1638
Db	5771	ttcttcaactcggggTAAAGAGACATCAATGAGATTGGAGATTACTTCACTGTGATAGA	5830
QY	1638	ngluuengluclnalaglnvalgluleuylscysgluvalglunhlsleuMetlysSerMe	1658

```

Db 5831 GGACCCGTGACAGA-----AAAGTTGAAGCTTGTCTAATGCAAAAT 5869
QY 1658 tliegluserlyserSerleuGluserleuGlnHisgluYHisAspThrleuGluGln 1678
Db 5870 GAAGAATTAAGACTCAAACTC-----CAATTACAGAGGTGACA 5908
QY 1678 nleuLeuAla-----LeuLYSGlnGlnMetGlnValValThrGln 1691
Db 5909 ACTAAATGACCAAAATTAAGACATGATAGAAATTTGGAATAAATAGTTGGGCACTTAAGAA 5968
QY 1691 ngluLYSGlnleuGlnGlnThrHisgluHisleuThrAlaGluValAspHisleuLY 1711
Db 5969 AGAAACTCAGATTTAAGTGAAGAAATTTGGAATTTCTTGTGATCAGACAGAGGATTACT 6028
QY 1711 sgluAsnIleGluLeu-----GlyLeuAsnPhelyAsnGluAlaGln-----GlnLYSTh 1728
Db 6029 CCAGAGAGTGAAGAACTTCTGCAAGGCTCAATTTGATTTAGAAATGCAATGCAATTAATC 6088
QY 1728 rThrLYSGlnGlnCysleuLeuAsnGlnLYS----- 1739
Db 6089 ATCAGCGTAAGATATTGAGATTAATGAGCCAGCAAGTGAATGACAGCTGGAAGAGAGATT 6148
QY 1740 ----GluLeuGlnGlnSerGlnHisArgleuGlnCysGln----- 1751
Db 6149 TCTTGATGTGGAAGAAATGAGCTGAGTAGATCAGATCGAAGAAAGCTAGCATGAGCATGA 6208
QY 1752 -----IleGluGlnleuMetLYSserLeuLYSAspLYSGlnSerAlaGlnGluThr 1768
Db 6209 AGCCCTCTACTGAGGCTGACTTAGAGTAGTTCACAAACAGACAGACTATGTTAGAAAA 6268
QY 1768 rleuLYSGlnSerGlnGlnLYSValIle--AsnleuAsnGlnGlnMetGlnMetLYSValMe 1787
Db 6269 AGACAATGAAGAAATTAACAGAGAGTATTGTCTGCTGAAGAAAGAACTCTCAGTGGTAC 6328
QY 1787 tleuGlnMetGlnGlnleuLYSAsnSerGlnArgThrValIleAlaGlnArg----- 1804
Db 6329 AAGTGAAGAGAAACAGCTTCTGTGAGAAATTAGATCTATCTCAAAAAAACAAGCGGCACT 6388
QY 1805 -AspGlnleuGlnAspAspLeuArgGlnSerValGlnMetSerIleGluThr----- 1821
Db 6389 GGAATCAGTTGTCTGAAGAAATTAAGAGAGAAACACAA--GAGCTTGAGCTCATCAAG 6445
QY 1822 -----GlnAspAspLeuArgLYSAlaGlnGlnAla 1832
Db 6446 TGAGTGTCTCCATTGCAATTCAGGTGCGAGGCGAGGCTGAAGAAAGCAAGCTCT 6505
QY 1832 uGlnGlnGlnLYSAspLYSValGlnGlnleuThrSerGlnIleSerValleuGlnGlu 1852
Db 6506 TCAGACTTTTCTCTGTGATGTGAGTGAAGCTTTAAAGACAAATCATCTCCAGGAAAA 6565
QY 1852 sIleSerleuLeuGlnAsnGlnMetleuLYSAsnValAlaThrValLYSGluThrLeu 1872
Db 6566 GCTGCAAGAGTTTGGAAAGAGACTCACAGGCACTGCTTTGACAAATATGAGCTGGAAGA 6625
QY 1872 rGluArgAspAspLeuAsnGlnSerLYSglnHisleuPheserGluIleGluThrLeu 1892
Db 6626 CCAAAATGCAACACTGAAATTAAGAGAAATTTGCTTCAGAAATCTCAAAAGCCCTGCA 6685
QY 1892 rleuSerleuLYSGluLYSGluPhe-----AlaLeuGlnGlnAla 1905
Db 6686 GGCCAGACTGATGATGCAATTAATGAAAAAGCTGAATGTCTCCAGAGCCTTGGAGGCCGC 6745
QY 1905 agluLYSAspLYSAlaAspAlaAlaArgLYSThrIleAspIleThrGluLYSleSerAs 1925
Db 6746 ACTGCGGAGAGAAAGTGAAGTTCATTGAGGCTGAGCTCAACAGAGAGAGAAATGCAATCA 6805
QY 1925 nIle-----GluGlnGlnleu 1931
Db 6806 GCTGAGAGAGAGCATGAGAAACTGAGAGTTCGATTCAGAGGCCGATGAGAAAGAGAGCAT 6865
QY 1931 uGlnGlnAlaThrAsnleuLYSGluThrleuLYSArgGlnSerleuIleGlnCys 1951

```

```

Db 6866 GCACATCGCAGAGAAACTGAAAAACGCGAGCGGAGAAATGATTACTT-----AA 6916
QY 1951 sgluGlnleuAlaLeuAsnThrGlnHisleuArgGluThrleuLYSserLYSAspLeuAla 1971
Db 6917 GCATTAAGTTGAGAACCTTGAAGAGAAATTTGCGAGATGTCAAGAAAGAACAGAGGCTACT 6976
QY 1971 aleuGlnLYSMetGlnGlnArgAspGluAlaAlaAsnLYSValIleAlaLeuThrGln 1991
Db 6977 GATT-----CTTGATGCGGAGAAATTCAAAGCA----- 7004
QY 1991 uLYSMetSerSerleuGlnGlnGlnIleAsnGlnValThrThrleuLYSGluLYS 2011
Db 7005 -GAATGAGACCTTAACAAACACAAATTAAGAGAGATGCCAGAGGCTGAAGATTTTGA 7063
QY 2011 u-----GlyGluLYSGlnThrPheLYSleuGlnArgProSerLY 2024
Db 7064 ATTAGACCTTGTCACGTTAAGTCTGAAAGAAAGAAATTCGACAAACAAATACAAAGAAA 7123
QY 2024 sgluGlnSerSerSerGlnMetGlnleuArgGlnSerleuLYSThrLYSAspLeuGln 2044
Db 7124 ACAAGGTCAGTTGTCTGAGAACTAGACAAAGTTACTCTCTCATTTTAAAGTCTGTAGAGAA 7183
QY 2044 nleuGlnGlnAlaGlnLYSGluLYSserGlnAlaThr----- 2056
Db 7184 AAGAGCAAGCAGAGATACAGATCAAGAGATCTAAAACTGCAGTGCAGATGCTTCA 7243
QY 2057 -AsnGlnIleLYSAsnleuThrAlaLYSleSerSerleu-----GluGlnGlnIle 2074
Db 7244 GATTCAGTTTAAGACTTAATGAGCAGATGACAGGCTTGTGGTGAACAAAGAAATTTAT 7303
QY 2074 uGln---AsnAlaSerIleleuAsnGlnAlaValSerGlnArgGlnAsnleuArgHis 2093
Db 7304 GAAGGCCACAGAAACAGAGTCTAGACCCACCACCAATAGAGAAAGACATCAGCTGGAATAG 7363
QY 2093 rLYSGlnGlnleuValSerGlnleuGln-----GlnLeuSerleuThr 2108
Db 7364 CATTGAAAGCTAGAGCCCGCTAGAGCTGATGAAAGAGACACTGTGTCTTACA 7423
QY 2108 uLYSserArgAspHisAlaPheAlaGlnSerLYSArgGlnLYSAspGlnAlaValAsnLY 2128
Db 7424 ACAACTGAAG-----GAAAGTGAATCATGACAGATTTACTTAAGGCTAG 7468
QY 2128 sIleAlaSerleuAlaGlnGlnIleLYSleuThrLYSleuThrLYSGlnMetAsp----- 2144
Db 7469 AGTGAAGAACCTTGAAGAGAGCTAGAGATGAGCAGACAAACCAAGAGATGCACCTCT 7528
QY 2145 -GluPheArgAspSerLYS-----GluSerleuGlnGlnSerSerHisleu 2161
Db 7529 TGAGCGCAGAGAAATTCACAAAGAGAGATGAGACCTTAAAGCAAAATTAAGAGGATGAC 7588
QY 2161 rGlnGlnleuCysThrLYSThrGlnleuGlnMetleuLYSGlnGlnLYSAspIle 2181
Db 7589 CCAAACTCTGAGAGCTGTGAATTAATGATGTTTACTATATGAGTCAAGAAAGAAATATCT 7648
QY 2181 eAsnAsnLYSleuAlaGlnLYSValLYSGlnValAlaArgGlnleuGlnHisleuSer 2201
Db 7649 GACAAATGATTTCAAAAGAGCAAGAGCAAGAAATATGTAATTA----- 7691
QY 2201 rleuLYSGlnGlnleuAspGlnIleGlnMetGlnleuArgAsnGlnLYSleuArgAsn 2221
Db 7692 -----GAAATTAATTAATTCATTCATTTGAATAT----- 7718
QY 2221 rGluLeuCysGlnLYSMetAspIleMetGlnLYSGlnIleSerValleuArgleuMetGln 2241
Db 7719 -----ATTTCACAAAGAAAAA 7735
QY 2241 nAsnGlnProGlnGlnGlnGlnAspAspValAlaGluArgMetAspIleleuGlnSer 2261
Db 7736 GCAAGAGAAAGATACAGATGAAGAAATTAACAACTGCCATGGAATGCTTCAACAA-- 7793
QY 2261 gAsnGlnGlnIleGlnGlnleuMetGlnLYSleSerAlaValLYSserGlnGlnHis 2281
Db 7794 -----CAATTAAAGAGCTCAATGAGAGAGAGGCGACCTGCAATTAATGACCAAGAAC 7846

```

Oy	2261	rLeuLeuSer-----	-----SerLeuSerSerGluLeuGlnLysGluThrGluAlaHisLys	2236
		:::		
Db	7847	CTGTAAAGGCCAAAGACAGCATCTGTAGTGCACAGTAGAG	-----	7886
Oy	2298	sHisCyMetLeuLeuHisnLeLysGluSerLeuSerSerThrLeuSerArgSerPheGlySe	2318	
Db	7887	-----TGCTT-----	-----GAACTTGAGAAAGGCTCAGTTGCTACAAAGGCGCTTGATGA	7930
Oy	2318	rLeuGlnThrGluHisValLysLeuHisnThrGlnLeuGlnThrLeuLeuAsnLysPheLys	2338	
Db	7931	GGCCAAAATAATATATATGTCTTTGGCAACCTTCAGTGAAGGCGCTCATTCAGAAAG	-----	7985
Oy	2338	sValValTyrArgThrAlaAlaValLysGluAspHisSerLeuLysAspTyrGlyLys	2358	
Db	7966	-----	-----GTAAAGAGTGGCAACCA	8002
Oy	2358	sAspLeuAlaLeGluGlnLysArgHisAspGluLeuArgLeuGlnLeuGlnCysLeuGln	2378	
Db	8003	GAACTGGACAGACAGATAGATACAAATTCAGTAGACTCGAAATAATTCAAATTCAGACACAGA	8052	
Oy	2378	uGlnHisGlyArgLysTyrSerAspSerAlaSerGluGlnLeuLysPhe-----	2394	
Db	8063	GCACCTGTGCTTAAACTGTGCCACAGGTGGAGAGAGACACCACTTGGAAAGACAAAG	8122	
Oy	2395	-CysGluLeGluPheLeuHisnLeuLeuPheLysLysAlaHisnLeileGlnSerVal	2414	
Db	8123	CTTGAAGATGCAAAATTCAGCGGTGGAAATTCGACGAAAGATCCAAAGCTCTCAATCCAA	8182	
Oy	2414	IGlnAspAspPheSerGlyValGlnValPheLeuHisnGlnValGlySerThrLeuGlnGln	2434	
Db	8183	AAATGCTCTTTCAGAGACACATTAAGAAGCTGCAGAGTCTTACAGAAATCTAGAGAG	8242	
Oy	2434	uGlnLeuGln-----HisLysLysLysPheMetGlnThrPheGlnGluLysPheGly	2450	
Db	8243	TGACCTTGAAATTCACAAAATTCGACAAAATGTGCTTGTGTAAGAAAGTAAACAAAATGAC	8302	
Oy	2451	-----	-----AspLeuHisValAspAlaLysLys-----Le	2459
Db	8303	TGCAAAGAAACCTGAGCTGCAGAGGGAATTCGATGAGTGCACAGAAACACAGACAGCT	8352	
Oy	2459	uSerGlnGlyMetGlnGlnGlnLysnArgArgLeaSerThrIleGlnLeuLeuThrLys	2479	
Db	8363	GCAAGGAAGCACTAGTGGAGAGGAAATTAAGCTAGCTGAGAGAGTTCGACGTTACTGTTGGA	8422	
Oy	2479	sArgLeuLysAlaValValGlnSerLysIleGlnArgGluIleThrValTyrLeuAsnGln	2499	
Db	8423	AGAAATTAAGAC-----	-----AGCAAAAGATCAA-----	8447
Oy	2499	nPheGluAlaLysLeuGlnGlnLysLysGlnGlnHisnLysGluLeuMetArgArgMetGln	2519	
Db	8448	-----	-----TTGAGAGAGCTGCACACTACAGAAATATGTAATTCAGAGAGAG	8466
Oy	2519	uHisHisGlyProSerAlaSerValMetGlnGlnGlnLysnAlaArgLeuLeuGlyLysLe	2539	
Db	8487	-----	-----AGCTAGATTGCTGCACAAAGCCAGGTGAGAAAGGAGAGG-----	8528
Oy	2539	uLysThrValGlnAspGlnLysSerLysLysLeuGlnSerArgIleLysMetLeuGlnAsnGln	2559	
Db	8539	-----	-----AAAGTAGAGAGGAATTAAGCTAAATATACGCTACGG-----	8564
Oy	2559	uLeuAsnLeuValLysAspAspAlaMetThrLysLysGlnLysLys-----ValAlaIleLe	2577	
Db	8565	-----	-----CTTCATGAGCTGAAAAGAAACACACAGCGCTTGTCT	8599
Oy	2577	uGlnAspLysLeuLeuSerArgAsnAlaGluAlaGluLeuLeuHisnAlaMetGlnValLysLe	2597	
Db	8600	TTTGTGAC-----	-----ACAAACAAACAGATATAGATAGAAATCCAGACATATCCAGACAGAAATT	8653
Oy	2597	uThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGlnIleGlnAsnLeuGlnLysMet	2617	
Db	8654	GACTTCTTAAAGAAAGATGTCTCAGTTCCACAGAAAGCTGGAGATAGACCTTTTAAAGTCTAG	8713	

Qy	2617	tValAlaLysGclYalAvalProTYrLysGluLileAspAen	-----	26313
Db	8714	T-----	-----AAGAAGAGCTCAAAATTCATTGAAAGCTACTAC	8749
Qy	2632	----LeuYsThrYsValValLysIleGluMetGluYsIleYsTYr-----	-----	26466
Db	8750	TCAGATTTTGGACAGATTTGAAGAAACCAAGATGACAACTCTAAATATATGTAATCACTT	88039	
Qy	2647	-----SerYsAlaThrAspGlnLileAlaTYrLysSerCys----	-----	26606
Db	8810	GAAAGAGAAATGATACGTGCCCGAGGAGAAATGAAAGTTGTTGATCAATTCCTGTAAACA	8869	
Qy	2661	-LeuGlnAspYsGclYsGluLileAlaGlyLeuYsGluLileAlaArgAlaGlnAl	26880	
Db	8870	GCTGAGAAAGGAAAAAGAGATTCGACGAAGAAACTCTCTCAACTTCACAGCTGCACAGA	89399	
Qy	2680	AspAspAspThrThrValCysValProLYsAspTYrGlnLYsAlaSerThrPheProVa	27006	
Db	8930	GAGCAGAAAAACA-----	-----89424	
Qy	2700	lThrCysGlyLysGlySerGlyLileValGlnSerThrAlaMetLeuValleGlnSerGcl	27206	
Db	8943	-----GGTACTGTTAGTATACCAGAGCTGATTAACAACGTA	89838	
Qy	2720	U-----LYsAlaAlaLeuGlu-----ArgGluLeuSerHisTYrLYsLY	27333	
Db	8984	GATCAAGAACTGAAAGAAACTCTGTGAAGAAAAACCAAGAGAGCATGAATTACTTGG	90433	
Qy	2733	AluSTYrHlSHlsLeuSerArgThrHmeSerSerSerGlnAspArgLYsThrLYsAl	27535	
Db	9044	TAAAGTACTGTTCTCTG-----CTTAAAGCCATGAAGAGTTAGAGAAAGCTTAAAGA	90934	
Qy	2753	AluSerAspAlaHisSerSerHisThrGlySerSerHis-----ArgGcl	27688	
Db	9095	GATGTACAGACACAAAGTGGCCCTCTGTCTACAGACATCTAAACAAGATTCCCGAG	91544	
Qy	2768	YsErProHlsYsThrGluThrLYrArgHlSHlsGlyProValThrProGluArgSerGluMe	27888	
Db	9155	GTCTCT-----TTGTAAGTCCAGTTGTTCCAGGCATCTCAAT	91966	
Qy	2788	TrProSerLeu-----HisLeuGlySerProLYsLeuSe	27999	
Db	9197	CCCTTCTTTACTGAAAAAGAGTTATCATCTGGCCAAATTAAGCTTCAGGCACAAAGCA	92566	
Qy	2799	rGluSerSer-----ThrLYsArgValAlaSerProAspArgSerGluLileTY	2815	
Db	9257	AAGATCCAGTGGAATATGGGAGATGTGAAGACCAACCCGTACCCAGAGAGCTT	9316	
Qy	2815	rSer-----GlnLeuValMetSer-----ProGly-LysThrGlyMetH	2828	
Db	9317	TTCTTAAAAAAGCAAGAAAGACAGTCAAGATGATTAATCCACCTTCAGAGAGACA-----	9369	
Qy	2828	lSHlsHlsIleLeuSerProSerLYsValLileuHlsLYs-----LYsArgAlaLeuS	28466	
Db	9370	-----CGAAGGTACTAGTTGTGGCCAGAGGAGACTTCCAGAAAGTTGTAAAGAAAGGTTTGC	9427	
Qy	2846	erProAspArg-----SerGluMetProHlArgHlsHiv	2857	
Db	9428	TGACATCCCGACAGAAAGACTAGCCATATCTCCGCAAGAACACAAACCATGSCAA-----	94833	
Qy	2857	AlleSerProGlyLYsThrGlyLeuHlsLYsAsn-----LeuThrG	28711	
Db	9484	-----CTTCGACGCCCGCCGCTGGCTGCACAGAGTTAGCGGTATCCCACTAGCTCTCG	95414	
Qy	2871	lUsErThrLeuPheAspAsnLeuSerSerProCysLYsGlnGlnLYsValGlnGluAsnL	28914	
Db	9542	CAAGAAATATC-----TTGAGAGTCTTCAAAACCAACAGCTGTGGCA-----G	9586	
Qy	2891	euaAsnSerProLYsGlyLYsIleLeuPheAsp-----ValLYsSerL	29044	
Db	9587	CAGATCAACAAAGGTCAAAGTTGCTCAGCGGAGCCAGTAGATTCAGGACCATCTCTCG	9646	
Qy	2904	YsSerMetProTYrCysProSerGln-----PhePheAspAsnSerLYsLeuGlyA	29211	

```

Db 9647 AGAACCCACACGAAATCCCTCCAGTCATTAATCTCTGAGAGAGTCCGACTACAG 9706
QY 2921 sphserGluLeuAaThrLagLUserAsnAspLysSerGlnAlaGluAsnTrp 2939
Db 9707 CCCCAGAGAGGCGCTGAGGCTCAAGCAGAGCCACCTTGTCCACCCCAAGCTGG 9762

RESULT 3
US-08-328-254-5
; Sequence 5, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,239
; FILING DATE: 22-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; TELEPHONE: (619) 535-9001
; TELECOMMUNICATION INFORMATION:
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 544..7990
; US-08-328-254-5

Alignment Scores:
Pred. No.: 7.6e-76 Length: 8789
Score: 1061.00 Matches: 625
Percent Similarity: 39.11% Conservative: 488
Best Local Similarity: 21.96% Mismatches: 936
Query Match: 7.18% Indels: 798
DB: 1 Gaps: 118

US-09-150-867-1 (1-2954) x US-08-328-254-5 (1-8789)
QY 529 LysGluAsnGlnLeuGlnTyrLeuProLysAspSerGlyAspMetAlaGluCysArgLys 548
Db 460 AAAAGAAAGAAATATGAAGATTG---AAAGAGAGAAAACTCTGTTTCTTGTGGAAA 516
QY 549 AlaserpheGluLysGluIleThrSerLeuGln-----GlnGlnLeuGlnSer--- 564
Db 517 AGTGAAGAAAGAAACTTTTACTCAGATGCAATCAGAAAGAAAACTCTCAGAGTAA 576
QY 565 -----LysGluGlnGluLysGln 571

```

```

Db 577 ATTATCACTTGGAAACTTGTCTGAAGACAGCAAAATATAAAGTCATGATACAGAG 636
QY 572 LeuValGlnSerPheGluLeuLysIleAlaGluLeuGlnLeuSerValLysAla 591
Db 637 AGAGTAAGAAACCTGAGAGATGAGAC-----AGAGAAACCTTAAGTCTCAGATC 684
QY 592 LysAsnLeuGluMetValThrAsnSerArgGluHisSerIleAsnAlaGluValGlnThr 611
Db 685 AGAAACCTTCACACAGCTGTAGACAGTAG-----TAGGTGAGGTAGAGCCAGAAA 738
QY 612 AspValGluLysGluValValArgLysGluMetSerValLeuGlyAspSerGlyTyr--- 630
Db 739 CTAGCTTATATGAGAGCTACAGCAAAA-----GCTGAGTTCTCAGATCAGAAACATCAG 792
QY 631 ---AsnAlaSerAsnSerLysSerGlnAspSerSerValAspGlyLysArgLysSerSer 649
Db 793 AAGGAATAGAAATATGTTGTTGAAGACTCTCAGCTTACTGGGCA----- 840
QY 650 SerHisAspGluCysIleGluHisArg---LysMetLeuGluGlnLysIleValAspLeu 668
Db 841 -----GTTGAAGATATGACACCAAGCTTCACTCTCAATATGAATATATGACAAA 894
QY 669 GluGluPheIleGluAsnLeuAsnLysLysSerGluAsn-----AspLysGlnLysSer 686
Db 895 GACCGGTGTACCAACACTTCATGCCAATATGAGAGCGCTCAGGAGATCTGTAAATATCC 954
QY 687 SerGluGlnAspPheMetGluSerIleGlnLeuCysGluAlaIleMetAlaGluLysAla 706
Db 955 AAGATGCTTCTCTGTGTCACAAATGAAGATCATCAGAGAAAGCTTTTG----- 1002
QY 707 AsnAlaLeuGluGluLeuAlaLeuMetArgAspAsnPheAspAsnIleLeuGlu--- 725
Db 1003 ---GCTTTGATCAGCAGCGCTGCCATCATTCCTTTGCAAAATATATGAGAACAA 1059
QY 725 ----- 725
Db 1060 GGAAGATGCGCTTCAGAGAGAGTGAATGCTTTTGAAAGCAGCAAAAGTCCGAAAT 1119
QY 726 AsnGluThrLeuLysArgGluIleAlaAspLeuLysArgSerLeuLysGluAsnGlnGlu 745
Db 1120 TCTGCCATCTTACAAATATAGAGTGAATTCATCTGAATTTTCAATAGCTCAAAAACAG 1179
QY 746 ThrAsnGluPheGluIleLeuGluLysGluThrGlnLysGluHisGluAlaGlnLeuIle 765
Db 1180 ATGAAC-----TCAGACCTGCAAAAGCAGTGA---GAGTGGTG 1218
QY 766 HisGluIleGlySerLeuLysLeuValGluAsnAlaGluMetLysAsnGlnAsnLeu 785
Db 1219 CAATCAAGAGAGAAATATAGAGAAATCTCATGAAAGCAGAAACAGATCAAACTTTT 1278
QY 786 GluGluAspLeuGluThrLysThrLysLeuLeuLysGluGlnGluIleGlnLeuAlaGlu 805
Db 1279 GTG-----GCTGAACCAAGTCAAGCGCATATGATGAGTACAGAAACACTTCTGCTCAC 1332
QY 806 LeuArgLysArgLysAspAsn-----LeuGlnLysValArgAsnPheAspLeu 822
Db 1333 CAGAAATGTTCTTGTGAACCTTAAGTCCCTTGAGAACAGAAAGAAAGCTGCAACTT 1392
QY 823 SerValSerMetGlyAspSerGluLysLeuCysGluGluIlePheGlnLeuLysGlnSer 842
Db 1393 TTAAATATATAGATAGAGTAAGTGAAGC-----GCAAGATTCACAAATATATAAAGAGC 1446
QY 843 LeuSerAspAlaGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuAspSer 862
Db 1447 -----AACCATCTACTTGAAGACTCTCTAAAGAGAGATACAACTTTTA---TTC 1491
QY 863 GluAsnLeuGluLeuLysGluLysMetGlyAspThrSerAsnTrpLysAsnGlnLysGlu 882
Db 1492 GAAACCTTAAGCTTGGAAGAAAGAAATAGATGTCATCTTCTCTAATATAAAGGAA 1551
QY 883 LysAlaAlaSerLeuPheGluLysGlnLeuGluThrGluLysSerAsnTrpLysLysMet 902

```

Db 1552 -----ATTGAGAGCTGACCCAGAGAAATGGACTTTAAGAAATT 1593
 QY 903 Glu1AspLeuGlnLysGlu-----LeuGlnSerAlaPheAsnGluLeu 918
 Db 1594 AATGATCCTTAAATCAAGAGAGATGAACTTAATCCAGAAAAGTAGAGTTTGCACAC 1653
 QY 919 TyrLeuAsnGlyLeuLeuAlaGlyLysValProArgAspLeuSerArgValGluLeu 938
 Db 1654 TATATAGAT-----GAAAG 1668
 QY 939 GluLysLysValSerGluPheSerLysGlnLeuGlu----- 950
 Db 1669 GAGAAAGCATTTGAGAGTTATCTGATCAGTACAGCAAGAAAACCTATTATTACTACAA 1728
 QY 951 -----LysAlaLeuGlu 954
 Db 1729 AGATGTGAAGAAACCGAAATGCATATGAGATCTTACTCAAAATACAAAGCACACAG 1788
 QY 955 GluLysAsnAla-----LeuGlnAsnGluValThrCysLeuSerGluTyrLys 970
 Db 1789 GAAAGAAATTTCTAAATTAGAATGCTTGCTAAATGAAATGCACCTAGCTTTGGAAAATAG 1848
 QY 971 PheLeuProAsnGluValGluCysLeuLysAsnGlnIleSerLysAlaSerGluGlu 990
 Db 1849 -----AAAAATGAGTTGGAACAGCTAAAGAACGCAATTTGCAAGAACACCAAGAAATTC 1902
 QY 991 MetLeu---LeuLysGlnGluGlyGlnHisSerAlaSerIleIleSerLysGlnGlu 1009
 Db 1903 TTAACAAATTAATGATTTGCTGAAAGAAAGAAATCAGAAATGATGCTGAGTTGAGACA 1962
 QY 1010 IleMetGln-----GluGlnSerLys 1016
 Db 1963 GTGAGAGACAGCTGAGATCTGAGATGACATACCAAAACAAATTCAGAGCGACAGCT 2022
 QY 1017 -----GlnIleLeuGlnLeuThrAspGluValThrHisThrGlnSerLys 1031
 Db 2023 GGTGGTTTAAACAAAGAAATCATGCTTAAAGGAAGAAACAAATTCGCAAAAGGAA 2082
 QY 1032 Val-----GlnGlnThrGlnGluGlnIlePheGluMetLysLysMetHisAsp 1047
 Db 2083 GTTAATGACTTATTACAAAGAAATGACAGCTGATGAAGGTAAATGAACTTAACATGAA 2142
 QY 1048 -----AspLeuPheGluLysLysThrLysArgAsnLysSerGlnAlaGlnAspLeuArg 1065
 Db 2143 TGTCAAAATCTAGAAATCGAAACCAATTAGAACTCTGAAAAGAAAGAGACAGACAGA 2202
 QY 1066 GluMetGluAsnLeuLysGlyThrMetGlu----- 1075
 Db 2203 AATCAATGTATTTTAAACCTCAGATGATCTTGAAAGTTAAAGAAATTTCTAGATAGT 2262
 QY 1076 -----SerValGluValLysIleAlaSerPheLysIleGluLeuGlu 1089
 Db 2263 TATAATGCCAGTTGGTGCACAAATAGAACCTATGCTAAGAAATTAAGAAATTTAACTAG 2322
 QY 1090 GluThrIleArgAspLysGluGlnLeuLeuHisGluLysLysThrPhePheGlnAlaMet 1109
 Db 2223 GAAAGTGAGAAAGAGAGAGACTGCTGCAGCATGAA-----TTA 2361
 QY 1110 GlnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLysLeuValGlu 1129
 Db 2362 CAGCAATTT-----AGAGAGATCTTGAAACCCAGCAATTTGCAAGAC 2403
 QY 1130 GlyAsnSerGlnAspProIleGluIleAsnAspTyr----- 1141
 Db 2404 ATGCAGTCACACAGAAATTAAGTGGCCTTAAGACCTGTCAAAATAGATCCGAAAGAAAGTAT 2463
 QY 1142 -----HisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCys 1157
 Db 2464 ATTTCAGGCGCTCATGAGTTGTCAACAAGTCAAAAGCAATGCAACACTTCAGTGCCT 2523
 QY 1158 LeuGluThrGluArgAsnSerLeuLysGlnValIleAspLeuAsnThrGlnLeuGln 1177
 Db 2524 CTGCAAAACAAATGAACACAGCTGATGAG-----CTAGAGAAATATATGTGAA 2571

QY 1178 SerLeuGlnAlaGlnSerIleGlu-----LysSerAspLeu 1189
 Db 2572 ATACTGACAGGCTGAAAGATGATGAACTGTAAGTCAAGTAAATGATCAAGTCAAGT 2631
 QY 1190 GlnLysProLysGlnAspLeuGluGluGluValLysLeuLeuGlnGluMetGluLeu 1209
 Db 2632 ATCAGACGACACAGAAATATGCGACAGAGAGTAGGGAACCTACTAATGCACTTAAATTA 2691
 QY 1210 LeuLysGly----- 1212
 Db 2692 TTAATGATGACAGTGTCTTCTCCATGCTGATGAGTGAAGACATACAGAGAGTGA 2751
 QY 1212 ----- 1212
 Db 2752 TTGGTGCAACCAACAAATGACAGACCCCTGTCTTTGGCTCCATTTGGACAGAGTAAT 2811
 QY 1213 -----HisLeuThrAspSerGlnLeuSerIleGluLysLeuGlnLeuGlnAsnLeu 1229
 Db 2812 TCCTACGAGCACTTGACA-----TTGTACAGCAAAAGAAAGTTCAAATGCACCTTGCC 2862
 QY 1230 GluValThrGluLysLeuGlnThrLeuGlnGluMetLysAsnIleThrIleGluArg 1249
 Db 2863 GAATTCAGAGAGAAATTTCTTATCTTACAAAGTGAACACAAATTTTACATGATCAGCAC 2922
 QY 1250 AsnGluLeuGlnThrAsnPheGluAspLeuLysAlaGlnHisAspSerLeuLysGlnAsp 1269
 Db 2923 TGTCAATGAGCTCTTAATATGACAGCTGCAAGCTGATGTTGACTCATTAAG----- 2976
 QY 1270 LeuSerGluAsnIleGluGlnSerIleGluThrGlnAspGluLeuArgAlaLeuGlnGlu 1289
 Db 2977 ---GCCGAAATTTGGTGTCTGCA-----ACGAATCTGAGAAACCTTCAGGT 3021
 QY 1290 GluLeuArgGluGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSer 1309
 Db 3022 GACTTGCTGGAAGAGATCAGCTG-----GGCTTGGAAGAGGGCTGCTTCATCCCTG 3075
 QY 1310 ValGlyIleSerSerProAsnHisAspAlaValAlaAsnGlnGluValSerLeuGly 1329
 Db 3076 TCATCTCTTGTGTGCTGACAGCTGTAGCTTAC-----AGTTTGGGA 3120
 QY 1330 GluValAsnSerLeuGlnSerGluMetLeuArgGluArgAspGluLeuGlnThrSer 1349
 Db 3121 GAC---TCCTCTTTTACAGAGCTCTTTAGAA-----CAGACAGA 3159
 QY 1350 CysLysAlaLeuValSerGluLeuGlu---LeuLeuArgAlaHisValLysSerValGlu 1368
 Db 3160 GATATGCTCTTTGAGTAATTTAGAGGGCTGTTTCAGCAAAACAGTGCAGTATAGT 3219
 QY 1369 GlyGluAsnLeuGlnIleThrLysLysLeuAsnGlyLeuGluLysGluIleLeuGlyLys 1388
 Db 3220 -----GAAGTATTT---TGCACAGCTGTGCAAGGAGAAATCTGACCAAG 3261
 QY 1389 SerGluGluSer-----GluValIleLysSerMetLeuGluAsn 1401
 Db 3262 AAAGAAACCCCTTGGCCCGCAAGAGGTGTGAAGAGCTTGAATCCCTCTGTAGAGTG 3321
 QY 1402 LeuLysGluAspAsnAsnLysLeuLysGlnAlaGlnGluGluTyrSerSerLysGluAsn 1421
 Db 3322 TACCGGAGCTCCCTGAGAGAGCTGAGAGAAAGTGAAGGATTAATGAAAAAT 3381
 QY 1422 GlnPheSerLeuGluGluValPheSerGlySerGlnLysLeuValAspGlnIleGluVal 1441
 Db 3382 AAG-----GAAATTCAGAG 3396
 QY 1442 LeuLysAlaGlnLeuLysAlaAlaGlnGluArgLeuGlnLysAspArgAspTyrPhe 1461
 Db 3397 CTCGAGCAGTATTATTAAGTTCTGAAGAGCAGCTTGACTGCTTAAGAGACAGATATTG 3456
 QY 1462 GluLeuValGlnThrAlaAsnThrAsnLeuValGluGlyLysLeuGluThrProLeuGln 1481
 Db 3457 TCA-----GAA 3462

QY 1482 AlaAspHisGluGlnAspSerIleAspArgSerGluGlnMetGluIleLysValLeu 1501
 DB 3463 AATGAACAGCTGGACACAGAAAGCTGACAAAGCTGTGAGATGAGTCCAGAGTTGGCG 3522
 QY 1502 GlyGluLysLeuGlnLysGlnIleLysLeuGlu-----ArgLeuGlnLys 1517
 DB 3523 GCAGAAAGAAACACAGCAGAAACACTGTCACTTGAGCTGGAAGTAGCAGACTCCAGCA 3582
 QY 1518 GluLysLeuGluLeuSerAsnLysLeuGluIleLysGlnMetGluThrSerVal 1537
 DB 3583 CAAGGTCTGAGCTTAAGTCTCTCG-----TCT 3609
 QY 1538 LeuLeuLysAspAspLeuGlnGlnLysLeuGlnSerLeuLeuSerGluAsnIleLeu 1557
 DB 3610 TTGCTGGCATCGACACAGAGATGCTATTCAAGGC-----CCA 3648
 QY 1558 LysGluAsnIleAspThrThrLeuLysHisLysSerAspThrGlnAlaGlnLeuGlnLys 1577
 DB 3649 AATGAGACGCTGACATATCAAAAGAAACATATTCTCAGAAACTACAGAAAGAACCCAAAG 3708
 QY 1578 -----ThrGlnGlnLeuGlnLeu----- 1584
 DB 3709 CATGATGTCATCAGATTTGTGATAAAGATGCTCAGCAGAGCTCAATCTAGCATTGAG 3768
 QY 1585 -----AlaLysAsnLeuAlaIleAlaLysAspAsnLysProIleThrGlnLys 1602
 DB 3769 AAAATAACAGACTGTGCTGATTGAACCCACAGAGAGTCTCTGGGAGACAGTCCCA 3828
 QY 1603 GluThr-----SerAlaAspCysValHis 1610
 DB 3829 GATACCAATTATGAGCTCCAGGGGAGATMAAACCCAGGCTCTTCAGATTCATTTCT 3888
 QY 1611 ProLeu-----GluGln 1614
 DB 3889 GAATTGCATTTTCTGCTCTAATGCTTTGCTGATCTTGCATTTCTCGGGAGATCAGAA 3948
 QY 1615 LysIleLeuLeuLeuThrGlnGlnLysLeuHisLysThrAsnGlnGlnLysLeu 1634
 DB 3949 GATATTCATATTTCTCAACTCGGGTAAAGACATCAATGAATTTGAGATTCTT 4008
 QY 1635 HisGluLysAsnGlnLeuGlnGlnAlaGlnValGluLeuLysCysGlnValGlnHisLeu 1654
 DB 4009 CATGTGATAGACCGCTGACAGA-----AAAGTTGAAGTTTG 4047
 QY 1655 MetLysSerMetIleGluSerLysSerSerLeuGlnSerLeuGlnHisLysAsp 1674
 DB 4048 CTAAATGAATGAAGAATTAGACTCAAAACTC-----CATTTA 4086
 QY 1675 ThrGlnGlnLeuLeuAla-----LeuLysGlnGlnMetGln 1687
 DB 4087 CAGGAGGTACAACTAATGACCAAAATGACATGATGAAATGGAATAATGATTGGC 4146
 QY 1688 ValValIleThrGlnLysLysGlnLeuGlnIleThrHisGlnHisLeuThrAlaGlnVal 1707
 DB 4147 GAACTTAAGAAAGAAACTCAGATTTAAGTAAAAATTTGGAATTTCTTGAGTAC 4206
 QY 1708 AspHisLeuLysGluAsnIleGluLeu-----GlyLeuAsnProLysAsnGlnAlaGln 1725
 DB 4207 CAGAGATTACTCCAGAGAGTAGAAACTTCTGAAGGCTCAATTGCTGATTGAATGCTAT 4266
 QY 1726 ---GlnLysThrThrLysGlnGlnLysLeuLysAsnGlnLys----- 1739
 DB 4267 GCAGATTAATCATCACGAGATATTGTGAGATATGTGGCCAGAGTAATGACACTGG 4326
 QY 1740 -----GlnLeuGlnGlnSerGlnHisArgLeuGlnCysGlu----- 1751
 DB 4327 AAGGAGATTTCTTGATGTGGAAAAATGAGCTGAGTACGATCAGATCGGAGAAAGCTAGC 4386
 QY 1752 -----IleGlnGlnLeuMetLysSerLeuLysAspLysGluSer 1764
 DB 4387 ATTGAGCATGAAGCCCTGACTGAGGCTGACTTAGAGGTAGTTCAAAACAGAGAAAGCTA 4446
 QY 1765 AlaLeuGluThrLeuLysGlnSerGlnLysValIle---AsnLeuAsnGlnGlnMet 1783

DB 4447 TGTATTGAAAAAGACATGAAAAATAGCAGAGCTATTGTGCTGCTTGAAGAAAGCTC 4506
 QY 1784 GluMetValIleLeuGlnMetGlnGlnLysAsnSerGlnArgThrAlaIleAlaGln 1803
 DB 4507 TCAGTGTCTACAACTGACAGAAACCGCTGTGTGAGATAATTAGTACTATGTCAAAAAA 4566
 QY 1804 Arg-----AspGlnLeuGlnAspAspLeuArgLysSerValGluMetSerIleGlu 1820
 DB 4567 ACCACGGACGAGATCATATTGTGTGAAAAAATGAAGGAGAAACACAA---GAGCTTGAG 4623
 QY 1821 Thr-----GlnAspAspLeuArgLysAla 1828
 DB 4624 TCTCATCAAGTGAAGTGTCTCCATTGCTCAGTGGCAGGCGAGGTAAAGAAAG 4683
 QY 1829 GlnGlnAlaLeuGlnGlnGlnLysAspLysValGlnGluLeuThrSerGlnIleSerVal 1848
 DB 4684 ACGGAACTCTTCAGACTTTGTCCCTGTGATGAGTGAAGTAAAGACAAACATCAT 4743
 QY 1849 LeuGlnGlnLysIleSerLeuLeuGlnGlnMetLeuTyraAsnValAlaThrValLys 1868
 DB 4744 CTCGAGGAAAGCTCAGAGTTGGAAAAAGCAGCTCACAGGCACTGTCTTGAACAAATGT 4803
 QY 1869 GluThrLeuSerGlnArgAspAspLeuAsnGlnSerLysGlnHisLeuPheSerGluIle 1888
 DB 4804 GAGCTGGAACCAATTCACACACTGATTAAGAAAGAAATTCCTCTCAAGGAATCT 4863
 QY 1889 GluThrLeuSerLeuSerLeuLysGlnLysLysLys-----Ala 1901
 DB 4864 GAAAGCTGCAGCCAGCACTGATGATCAGATTTGAAAACTGATGCTCCAGGCC 4923
 QY 1902 LeuGlnGlnAlaGlnLysAspLysAlaAspAlaAlaArgLysThrIleAspIleThrGlu 1921
 DB 4924 TTGAGGCCGCACTGTGTGAGAAAGTGTGCAATTGAGCTGAGCTCAACACAGAG 4983
 QY 1922 LysIleSerAsnIle-----Glu 1927
 DB 4984 GAAGTGCATCAGCTGGAAGAGCATCGAAGAACTGAGATTCGCATTGAGGCCGATGAA 5043
 QY 1928 GluGlnLeuLeuGlnGlnAlaThrAsnLeuLysGluThrLeuThrGlnArgLysSerLeu 1947
 DB 5044 AAGAACACCTGTCACATTCACAGAAACTGAAGAAACCGCAGCGGAGATATTCACTT 5103
 QY 1948 IleGlnCysLysGlnGlnAlaLeuAsnThrGlnHisLysArgGluThrLeuLysSer 1967
 DB 5104 -----AAGATAAAGTTGAGACCTTGAAGGAAATTCAGATGTCAGAAACAAAC 5154
 QY 1968 LysAspLeuAlaLeuGlnLysMetGlnGlnLysArgAspGlnAlaAsnLysValIle 1987
 DB 5155 CAGGAGTATGTGATT-----CTTGATGCCGAGAAATTCCAAAGCA----- 5193
 QY 1988 AlaLeuThrGluLysMetSerSerLeuGlnGlnIleAsnGlnLysValIleThrThrLeu 2007
 DB 5194 -----GAGTAGAGACTTAAACAAACAAATGAGACAGATGGCCAGAAAGCTG 5241
 QY 2008 LysGlnGlnLysLys-----GlyLysGlnLysThrPheThrLeuGln 2020
 DB 5242 AAAGTTTAAATTAAGACCTTTCACAGTTAAGCTGCAAAAGAAATTCGACAAACAA 5301
 QY 2021 ArgProSerLysGlnGlnSerSerSerGlnMetGlnGlnLeuArgLysSerLeuLysThr 2040
 DB 5302 ATACACAAAAACAGAGTCAAGTTGTCAGAACTGAGAACAGTTACTCTTCATTAAAG 5361
 QY 2041 LysAspLeuGlnGlnGlnAlaGlnLysGlnIleSerGlnLysAlaThr 2056
 DB 5362 CTGTTAGAAAGAAAGGAGCAGAGATACAGATCAAGAAACAAATCTAAACTGCAAGT 5421
 QY 2057 -----AsnGlnIleLysAsnLeuThrAlaLysIleSerSerLeu-----Glu 2070
 DB 5422 GAGATGCTTCAGATATCAGTTAAAGAGCTAAAGAGCAGTACAGGCTGTGTGTGATG 5481
 QY 2071 GluGlnIleLeuGln---AsnAlaSerIleLeuAsnGlnLysAlaValSerGlnArgGluAsn 2089

D	5482	CAGAAATTTATGAAGGCCACAGAACAGAGCTTAGACCCACCAATTAGAGAGACGATCAG	5541
Q	2090	LeuATgHisSerLysGInGlnLeuValSerGluLeuGlu-----GlnLeu	2104
D	5542	CTGAGAAATATGACATTGAAAGAGCTGAGAGCCCGCTGAGAGCTGATGAAAGAGACGCTC	5601
Q	2105	SerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArgGluLysAspGlu	2124
D	5602	TGTGCTTTACAAACAACTGAGAG-----GAAAGTGACATCATCATGACATTTTA	5646
Q	2125	AlaValAsnLysIleLeaSerLeuAlaGluGlnIleLysIleLeuThrLysGluMetAsp	2144
D	5647	CTTAGGGGTAGAGAGGAGAACTTGAAAGAGAGCTAGAGATTCAGATGAGCCAGACAAACCAAG	5706
Q	2145	-----GlnPheArgAspSerLys-----GlnSerLeuGlnGlnLeuSer	2157
D	5707	CATGCACCTCTTGAGCGACAGAAATTTCCAAAGAGAGAGCTGAGAACCTTAAGCAAAAAATA	5766
Q	2158	SerHisLeuSerGluGlnLeuCysThrTyrrLysThrGlnLeuGlnMetLeuLysGlnGln	2177
D	5767	GAAAGGATGACCCAAAGCTGACAGAGCTGCGAATTCAGATTCGTTACTATATAGTCACAGA	5826
Q	2178	LysGlnAspIleAsnAsnLysLeuAlaGluLysValLysGluValAspGluLeuLeuGln	2197
D	5827	AAAGAAATCTGCAAAATGAAATTAACAAAAAGACCAAGACGGAATCTGAAATTA-----	5880
Q	2198	HisLeuSerSerLeuLysGlnGlnLeuAspGlnIleGlnMetGlnLeuArgAsnGluLys	2217
D	5881	-----GAAATATATAAATTCATCA	5988
Q	2218	LeuArgAsnTyrrGlnLeuCysGluLysMetAspIleMetGluLysGluIleSerValLeu	2237
D	5899	TTTGAAATAT-----ATTATG	5913
Q	2238	ArgLeuMetGlnAsnGlnUProGlnGlnGlnGlnLysAspValAlaGlnArgMetAspIle	2257
D	5914	CAAGAAAAAGACAGAAAGAAAGTACAGATGAAAGAAATCAACAGACCTCATAGAGATG	5973
Q	2258	LeuGlnSerArgAsnGlnGlnIleGlnGlnLeuMetGlnLysIleSerAlaValTyrSer	2277
D	5974	CTTCAAAACA-----CAATTAAAGAGCTCATGAGAGAGTGGACCCCTGCATATAT	6024
Q	2278	GlnGlnHisThrLeuLeuSerSerLeuSerSerGluLeuGlnLysGluThrGlnAlaHis	2297
D	6025	GACCA-----GAAGCCGT	6039
Q	2298	LysHisCysMetLeuAsnIleLysGlnLeuSerSerThrLeuSerArgSerPheLys	2317
D	6040	AAG-----GCCAAAGACAGAAATCTTAGTACATGACATGAG	6075
Q	2318	SerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPhe	2337
D	6076	TGCTTAAACTTGAGAAAGGCTCAGTTGCTACAAAGGCTTGATGAGAGCCAAAAATATATAT	6135
Q	2338	LysValValTyrrArgThrAlaAlaValLysGlnAspHisSerLeuIleLysAspTyrGlu	2357
D	6136	ATTGTTTTCGAATCTTCAGAG-----AATGCCCTCATTCAGAAAGTAGAA	6180
Q	2358	LysAspLeuAlaAlaGluGlnLysArgHisAspGlu-----LeuArgLeuGlnLeu	2374
D	6181	GATGGCAAGCAGAAACTGTGAGAAAGAGATGAAAGAAATCAGTAGACTGAAAAATCAAAAT	6240
Q	2375	GlnCysLeuGlnGlnHisGlyArgLysTrpSerAspSerAlaSerGlnGluLeuLysPhe	2394
D	6241	CAGAGCCAGAGAGCTGCTCTCTTAACCTGTCCAGGTGGGAAGAGACACCAACTTTGG	6300
Q	2395	-----CysGlnIleGluPheLeuAsnGlnLeuLeuPheLysAlaAsnIle	2410
D	6301	AAGAGCAAAACTTAGAAGTGAAGAAATCTGACAGAGTGAATTCGACACAGATCCAGAGT	6360
Q	2411	IleGlnSerValGlnAspAspPheSerGluValGlnValPheLeuAsnGlnValGlySer	2430
D	6361	CTCAAAATCCAAAAATGCTCTTTGACAGACACATTTAGAAGTGTCTGCAAGTCTTTCAAG	6420

QY	2431	ThrLeuGlnGlnGluLeuGlu---	HisLysLysGlyPheMetGlnTrpLeuGlnGluPhe	2449
Db	6421	AAATCTAGAGAAATGAGCTTGAATGATGACAAAATGAGCAAAATGCTCTTGGTCTTCAAAAACATA		6480
QY	2450	Gly-----	AspLeuHisValAspAlaLys	2458
Db	6461	AACAAAATGACTGCAAAAGAAACCTGAGCTCACAAGGAAATGCAATGATGATGCGACACAGAA		6540
QY	2459	-----LeuSerGlnGluMetGlnGlnLysAsnArgArgIleAlaSerThrIleGln		2475
Db	6541	ACAGACAGCTGCGCAAGAAACATCAGTCGAGAGAAAATAGCTACTGAGACAGTTGCAG		6600
QY	2476	LeuLeuTrpLysArgLeuLysAlaValValGlnSerLysIleGlnArgGluIleThrVal		2495
Db	6601	TTACTGTTGGAGAAATPAAAGAGC-----AGCAAAAGATCA-----		6636
QY	2496	TyrLeuAsnGlnPheGlnAlaLysLeuGlnGluLysGlnGlnAsnLysGluLeuMet		2515
Db	6637	-----TGGAAGCAGCTCACACTGAAAATAGTGAATTGAAG		6672
QY	2516	ArgArgMetGlnHisHisGlyProSerAlaSerValMetGlnGlnLysAsnAlaArgLeu		2535
Db	6673	AAG-----AGCTGATTTGATGATCACAAAGCACCGGTGGAAAG		6711
QY	2536	LeuGlyIleLeuLysThrValGlnAspGlnSerLysLeuGlnGlnSerArgIleLysMet		2555
Db	6712	GAAAGG-----AAAGTCAGAGAGAAATAGCTGAATTCACAGCTACG-----		6733
QY	2556	LeuGluAsnGluLeuAsnLeuValLysAspAspAlaMetHisLysGluLysLys----		2573
Db	6754	-----CTTCATGAAAGCTGTAAGAAACAAACAC		6777
QY	2574	ValAlaIleLeuGlnAspLysLeuLeuSerArgAsnAlaGluAlaGluLeuAsnAlaMet		2593
Db	6778	CAGGCTTTGCTTTGGAC-----ACAAACAAACAGTATGAAGTGAATTCACAGCATAC		6811
QY	2594	GlnValLysLeuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGluIleGluAsn		2613
Db	6832	CGAGAGAAATGACTCTTAAGAAGAAATGCTCAGTTCACAGAGAGTGGAGTATACCTT		6891
QY	2614	LeuGlnLysMetValAlaLysGlyAlaValProTyrLysGlnGluIleAspAsn----		2631
Db	6892	TTAAAGCTTACT-----AAAGAAAGCTCATATATTCATTC		6927
QY	2632	-----LeuLysThrLysValValLysIleGluMetGluLysIleLysTyr		2646
Db	6928	AAAGCTACTACTCAGATTTTGGAGAATTCAGAAAACCAAGTGTGCACATCTAAATAT		6987
QY	2647	-----SerLysAlaThrAspGlnGluIleAlaTyrLeuLys		2658
Db	6988	GTAATACAGTTGAAAGAAATGAACGCGCCAGGGAATGAATGAGTTGTTGATCAAA		7047
QY	2659	SerLys-----LeuGluAspLysGlnGlnGluLysLeuArgArgLeuLysGlnGluLeuArg		2676
Db	7048	TTCTGTAAACAGCTGCGAAGAGGAAAGAGAAATGATCTCAAAAGAACTCTTCACACTTCAA		7107
QY	2677	ArgAlaGlnAlaAspAsnAspThrThrValCysValProLysAspTyrGlnLysAlaSer		2696
Db	7108	GCTGCACAGGAGAGACAGAAAACA-----		7131
QY	2697	ThrPheProValThrCysGlyGlySerGlyIleValGlnSerThrAlaMetLeuVal		2716
Db	7132	-----GGTACTGTTATGATACCAAGGTGCATGAA		7161
QY	2717	LeuGlnSerGlu-----LysAlaAlaLeuGlu-----ArgGluLeuSer		2729
Db	7162	TTAAACAACCTGAGATCAAAAGAACAGAAACCTTGAAGAAAGAAAACCAAGAGGACAGAT		7221
QY	2730	HisTyrLysLysLysTyrHisHisLeuSerArgThrMetSerSerGlnLysAspArgLys		2749
Db	7222	GAAATCTTGGATGTAGTACTTCTTCCCTTG-----CTATTAAGCCAGCAAAAAGTTATGAG		7272

QY 301 AlalysThrValIleIleCysThrIleThrProValSer-----PheAspGluThrLeu 318
 DB 994 GCCAAAGACGATATATGCAAAATGTTTCATCTCGATGCTGGTGGGGAACCCCTTA 1053
 QY 319 SerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGlu 338
 DB 1054 TCACACTTAATCTTGGTCCCAAGAGCCAGATGATTAATAAACAAGAGCAGTAAATGAA 1113
 QY 339 ValLeuAspArgGluAlaLeuLeuLysArgTyrArgLysGluIleLeuAspLeuLys 358
 DB 1114 -----GACACCCCAAGGAAATGTGGAGCAGCTCCAGAGCTGAAGTGAAGAGCTCAAAAGAA 1167
 QY 359 GlnLeuGlnAsnLeuGlnLysSerSerGluThrLysAlaGlnAlaMetAlaLysGlnGlu 378
 DB 1168 CAGCTGGGAGAGCTTGGCTTCAGACAGACAGCCAGAAAGCTCTCCAGAGAGCAAA 1227
 QY 379 HisThrGlnLeuLeuAlaGluIleLysGlnLeuHisLysGluArgGluAspArgIleTyr 398
 DB 1228 AAGAAGACTAATATATGAGATATTCCAG-----GAAGCAATGTTATTC 1272
 QY 399 HisLeuThrAsnIleValAlaAlaSerSerGlnGlnLysGlnAspGlnArgValLys 418
 DB 1273 TTT-----AAGAAATCTGAACAGGAAGAAAG----- 1299
 QY 419 ArgLysArgArgValIleThrPAlaProGluLysIleGlnAsnSerLeuHisAlaSerGlu 438
 DB 1299 ----- 1299
 QY 439 ValSerAspPheAspMetLeuSerArgLysLeuProGluAsnPheSerLysLysAlaLysPhe 458
 DB 1300 ---TCTGTGATAGAAAGATTACCAATTTAGAGACCTCAGCTCCAGAAAGAAAGAAATTT 1356
 QY 459 SerAspMetProSerPheProGluIleAspAspSerValCysThrGlnPheSerAspPhe 478
 DB 1356 ----- 1356
 QY 479 AspAspAlaLeuSerMetMetAspSerAsnGlyIleAspAlaGluThrAsnLeuAlaSer 498
 DB 1357 -----ATTCATCTATATATAAATGATTGTGAAATTTCCGA---GAGCAT 1395
 QY 499 LysValThrHisArgGluLysThrSerLeuHisGlnSerMetIleAspPheGluGlnIle 518
 DB 1396 CAATATATACCTGTGAAAG-----CTCCAC----- 1422
 QY 519 SerAspSerValGlnPheHisAspSerSerLysGlnAsnGlnLeuGlnTyrLeuProLys 538
 DB 1423 -----AAGAAATCCCGGAGAGTTTCTGCTGCTGAG 1452
 QY 539 AspSerGlyAspMetAlaGluCysArgLysAlaSerPheGluLysGluIleThrSerLeu 558
 DB 1453 GAGCAGAGTGTGTTGCCTC-----TCAGAAATTAAGAGATGAGATTCAGAACTGTG 1500
 QY 559 GlnGlnGlnLeuGlnSerLysGlnGlnLysLysGlnLeuValGlnSerPheGlnLeu 578
 DB 1501 CGAAGAAATATAGAGCAACAGCCAGAGTTGCCAAAGTATGCTATGCAAAATCATTTCCCTC 1560
 QY 579 LysIleAlaGluLeuGlnGlnLeuSerValLysAlaLysAsnLeuGlnMetValThr 598
 DB 1561 AGG-----GAGGAGAAATAGAAAGACTGAGATTA-----TTAGAGCCTGTGAAA 1602
 QY 599 AsnSerArgGlnHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValAla 618
 DB 1603 AGAGCTCAAGAA----- 1614
 QY 619 ArgLysGlnMetSerValLeuGluAspSerGlyTyrAsnAlaSerAsnSerAspLeuGln 638
 DB 1614 ----- 1614
 QY 639 AspSerSerValAspGlyLysArgLysSerSerHisAspGluCysIleGlnHisArg 658
 DB 1614 ----- 1614

QY 659 LysMetLeuGlnGlnLysIleValAspLeuGlnGluPheIleGlnAsnLeuAsnLysLys 678
 DB 1615 ---ATGATGCCAGACACCATTCGAAACAACTAGAAAAGCTTTCTCGAATATAGTGGCATG 1671
 QY 679 SerGlnAsnAspLysGlnLysSerSerGlnGlnAspPheMetGluSerIleGlnLeuLys 698
 DB 1672 GAGAAATATGACAA-----AATCAGCAGAGATTTTCCACCTTAAGAGCTCAGAAAGAG 1722
 QY 699 GluAlaIleMetAlaGlnLysAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAspAsn 718
 DB 1723 CCATGTTTGTTCGCA----- 1737
 QY 719 PheAspAsnIleIleLeuGlnAsnGlnLeuThrLeuLysArgGluIleAlaAspLeuGluArg 738
 DB 1738 -----AACACTGAGAAAGTTAAAGACACACTCTCCAAATTCAGCA 1779
 QY 739 SerLeuLysGluAsn---GlnGluThrAsnGluPheGluIleLeuGluLysGluThrGln 757
 DB 1780 GAGCTGATTAATTCAAACAGAAATATGAGAAATTC-----AAGAACTTACT 1827
 QY 758 LysGlnHisGlnAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn 777
 DB 1828 AGGAAAGCAGCTAGAAATTCAGATCAGAGCTTCAGCTTTTGCAGAAAGCG----- 1878
 QY 778 AlaGluMetTyrAsnGlnAsnLeuGlnLysAspLeuGluThrLysThrLysLeuLeuLys 797
 DB 1879 -----AACCTTAATCTTCAAAAACCTTTGGAAAGCA---ACAAAAGCCTGCAG 1923
 QY 798 GlnGlnGlnIleGlnLeuAlaGlnLeuLysArgLys---ArgAlaAspAsnLeuGln----- 814
 DB 1924 CGGCAGAAAGTT-----TTCAGCTGAAATTAATTCGTGAAACACTTAAGATTA 1977
 QY 815 -----LysLysValArgAsnPhe-----AspLeuSerValSer 825
 DB 1978 ACTACACCAACCAAGCCCTTACCACCTTCAATCCGACCAAGTACCAAAATTAAGCCCTGAA 2037
 QY 826 MetGlyAspSerGluLysLeuCysGlnGluIlePheGlnLeuLysGlnSerLeuSerAsp 845
 DB 2038 ATGGGAACCTTGGCTGCTCTA-----TACACTCAGAAATTCAGACATATTAGAT 2085
 QY 846 AlaGlnAlaValIleThrArgAspAlaGlnLysGluCysSerPheLeuArgSerGluAsnLeu 865
 DB 2086 AATGATATATTAATATGACCCAGCTTCCTCTGAG----- 2118
 QY 866 GluLeuLysGluLysMetGluAspPheSerAsnTyrArgGlnLysGluValAla 885
 DB 2119 -----ATGATGAACCAAGCTTT 2136
 QY 886 SerLeuPheGluLysGlnLeuGluThrGluLysSerAsnTyrLysLysMetGluAlaAsp 905
 DB 2137 GAGGCCATTTCTGAAGACTTGAACAGTGAAGCAAGAAAGAGTGCCTTCAGGCCAAA 2196
 QY 906 LeuGlnLysGlu-----LeuGlnSerAlaPheAsnGluIleAsnTyr 919
 DB 2197 CTGATGTAAGAAGACATTAATAACCTTAAGCTTCAGCGCATGTGTGACAAACTGGAACAT 2256
 QY 920 LeuAsnGlyLeuLeuAlaGlyLysValProArgPheLeuSerArgValGluLeuGlu 939
 DB 2257 CATTCACCCCAATG-----CAGAGCTTTCTCTATCCGAAGAATTT--- 2298
 QY 940 LysLysValSerGluPheSerLysGlnLeuGlnLysAlaLeuGlnGluLysAsnAlaLeu 959
 DB 2299 -----GATTGGACCAAAACAGAGAGAGAGCTTCTCAGCACTTAAGTGTCTT 2346
 QY 960 GluAsnGluVal---ThrCysLeuSerGluTyrLysPheLeuProAsnGluValGluCys 978
 DB 2347 GAAAGCAGCTTCAAGACACTCAAACTAAATGACTTTTGAAGAGGAGCTACATGAC 2406
 QY 979 LeuLysAsnGlnIleSerLysIleSerGlnIleMetLeuLeuLysGlnGlu----- 996
 DB 2407 CTGGGAGTAGTCTTCATTTGCTGTGACAAAGAGCTTTCTTAGAGAAATTTGGAATTAAGT 2466
 QY 997 -----GlyLysHisSerAlaSerIleIleSerLysGlnGluIleMet 1011

```

Db 2467 TCATCAAAAAGATCAGAGAAAGATTCAACAAACTTCCGAAAGACATCGCATGA 2526
QY 1012 GlnGlnGlnSerLeuGlnIleLeuGlnLeuThrAspGluValThrHisThrGlnSerLys 1031
Db 2527 CACCTTCATTAATTAATATATCTCAGCTTGAAGAAAGAAAGCTCTGAGCAAAACCTGCG 2586
QY 1032 ValGlnGlnThrGlnGlnGlnIleThrLeuGlnMetLysLysMetHisAspAspLeuGlu 1051
Db 2587 CTCACAGATTCCTATGACAACTTACAGAAATATGAAATTTGACATGATGACCACTTCA 2646
QY 1052 LysTyrIleArgAsn---LysSerGluAlaGluAspLeuLeuArgGlnMetGlnLeu 1070
Db 2647 AGAAACCTCCAAAGCTTCAAAAGAAAGAAATGAACCTGAAATCTGATCTGAAATTAATTG 2706
QY 1071 LysGlyThrMetGlnSerValGlnValLys-----IleAla 1082
Db 2707 ATGAGACTTCCTGAGCGACAAAGAAAGCAATTAACAATTAATCACTTACAGTTGAAAGAA 2766
QY 1083 AspThrLysHisGlnLeuGlnGlnIleThrIleArgAspLysGlnGlnLeuHisGlnLys 1102
Db 2767 GATTAAGAAAGAAAGCTTCAAAAGAAATCTTAAGAAATCTTGAAGCGCTGATACGACAGAGAA 2826
QY 1103 ---LysTyrPhePheGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSer 1121
Db 2827 CAGAAAGAGACGCGCCAGAGTGTGAGCAGACAGATGCGCAAAAGTACAGAACTGAGAGAGAGC 2886
QY 1122 LeuProSerLysLeuValGlnLysAsnSerGlnAspProIleGluIleAsnAspTyr 1141
Db 2887 TTG----- 2889
QY 1142 HisAsnLeuIleAlaLeuAlaThrGlnLysArgAsnAsnIleMetValCysLeuGlnThrGlu 1161
Db 2890 -----CTTGGCTTCACTGAAAAA-----GTGATCACTTCCCTGGAAGAAAGTCT 2928
QY 1162 ArgAsnSerLysGlnGlnValIleAspLeuAsnThrGlnLeuGlnSerLeuGlnAla 1181
Db 2929 AGAGATTCGTGAATGAAGATTGTAGCTGACCTCATGAAACCGATCCAGAGACTTAAGTCA 2988
QY 1182 GlnSerIleGlnLysSerAsp---LeuGlnLysProLysGlnAspLeuGlnGlnLysGlu 1200
Db 2989 TCGGTCTGTGAAAGAAACGAAACTATGACACCTCGAAACAGAACGTAAGACATTAAT 3048
QY 1201 ValLys-----LeuLeuGlnMetGlnLeuLeuLeuLysGlnHisLeuThrAspSer 1217
Db 3049 TGCAAATTCACACTCTGCTTGTGTTGACAGAGAA-----GAGAGC 3087
QY 1218 GlnLeuSerIleGlnLysLeuGlnLeuGlnAsnLeuGlnValThrGlnLysLeuGlnThr 1237
Db 3088 AGAGTGTGATCAGAAAGAGAGAGATGATTTCTGATCTGAAAGAAACCTTAGG--- 3144
QY 1238 LeuGlnGlnGlnMetLysAsnIleThrIleGlnLysAsnGlnLeuGlnThrAsnPheGlu 1257
Db 3145 -----CTGAGAAATCACTTCTCTGAG 3162
QY 1258 AspLeuLysAlaGlnHisAspSerLysGlnAspLeuSerGlnAsnIleGlnGln--- 1276
Db 3163 GACATA-----GAGAGGATATGCTCTGTGAGACCTGCTCATGCCACTGACAGCTG 3216
QY 1277 -----SerIleGlnThrGlnAspGlnLeuArgAlaAlaGlnGlnGln 1290
Db 3217 AACATGCTCACAGAGCCTCAAAAAAACCTCGGGCTGCTGCACTCTCCAGAGAGAA 3276
QY 1291 LeuArgGlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuAspLysSerVal 1310
Db 3277 CTCACCAAGAGAGAGCCGTGATTCAGAACTTCACAGCAAGCTA----- 3321
QY 1311 GlyIleSerSerProAsnHisAspAlaValAlaAsnGlnGlnLysValSerLeuGlnGlu 1330
Db 3322 -----AACCAAAAGAAA----- 3333
QY 1331 ValAsnSerLeuGlnSerGlnMetLeuArgGlnArgAspGlnLeuGlnThrSerCys 1350
Db 1331 ValAsnSerLeuGlnSerGlnMetLeuArgGlnArgAspGlnLeuGlnThrSerCys 1350

Db 3334 -----GAGGAAGTAGAACAGAGAGAGATGAATTAATCACTCAAAATG 3375
QY 1351 LysAlaLeuValSerGlnLeuGlnLeuLeuArgAlaHisValLysSerValGlnGln 1370
Db 3376 AGGCAACTA-----GACATGTGATGCAATTCCTGCTGCTGAG 3411
QY 1371 Asn-----LeuGlnIleThrLysLysLeuAsn 1379
Db 3412 GATCCCAAGAGCTTAAAGACACACCTCACTTCAAAACATTTGGCAAAACCTCGGA 3471
QY 1380 GlyLeuGlnLysGlnIleLeuGlnLysSerGlnGlnSerGlnValLeuLysSerMetLeu 1399
Db 3472 ACACAAAGAAAGAGATA-----GAAGATGGAAGAGCCCTGTAAGACTTCTTG 3519
QY 1400 GlnAsnLeuLysGlnAspAsnLysLysLeuGln-----GlnAlaGlnLysTyrSer 1417
Db 3520 GAACACCTTGTAAACAAAGCTTAATGAAGACAGAGAAAGTCAAAATCTGAAATCCAGAA 3579
QY 1418 SerLysGlnAsnGlnPheSerLeuGlnGlnValPheSerGlySerGlnLysLeuValAsp 1437
Db 3580 ATGAGAGAGCAGTTGCGTGAAATGAAACCTAGCCCTGGAAGCTCAGCAGTTAAATAGAG 3639
QY 1438 GlnIleGlnValLeuLysAlaGlnLeuLysAlaAlaGlnGlnLysArgGlnLysAsp 1457
Db 3640 AAAAAGTGGCTCTGCAAGGTGAGTGTGATTTAAAGACAAAGAGAAACAGTGTAT 3699
QY 1458 ArgAspTyrPheGlnLeuValGlnThrAlaAsn----- 1468
Db 3700 CAGAAATCATCCAGATATATCAACACGCTGAAGAAATGAACAGAAAGTATCAAGAAAGA 3759
QY 1469 -----ThrAsnLeuValGlnGlnLysLeuGlnThrProLeuGlnAlaAspIleGlu 1485
Db 3760 CTGCAAAAGTAATTAATTTGAGAAATGCTGAAA-----ATGAAGACAGCTTAGAA 3813
QY 1486 GlnAspSerIleAspArgArgSerGlnGlnMetGlnIleLysValLeuGlnLysLeu 1505
Db 3814 GAAGTCCAAAGTGCCTTTACACAAAGAGATGATGCTTAAGATCATGATCATGATCAATC 3873
QY 1506 GlnArgAsnGlnTyrLeuLeuGlnGlnArgLeuGlnGlnGlnLysLeuGlnLeuSerAsnLys 1525
Db 3874 GAACGAACCCAAACTTGTGAGTCTTAAGCATTCACGAAAGAAAGCAACTAGATCAAG 3933
QY 1526 LeuGlnIleLeuGlnLysGlnMetGlnThrSerValLeuLysAspLeuGlnGln 1545
Db 3934 CTGGAAGAAATGTATGAAGAAAGAGAGACATCCAGAGATGAGAAATTTAAGAGAG 3993
QY 1546 LysLeuGlnSerLeuSerGlnAsn-----IleIleLeuLysGlnAsnIleAspThr 1563
Db 3994 CAGGTGAGTGTCTGCTGAGGAAATGAAAGTTGGTAGTCAACCAAAATTTG----- 4047
QY 1564 ThrLeuLysHisLysSerAspThrGlnAlaGlnLeuGlnLysThrGlnGlnGln 1583
Db 4048 -----CATCAAGATTCAGTACGTACGTACGACATTAAGAAAGAAAGTCAAGG 4095
QY 1584 LeuAlaLysAsnLeuAlaIleAlaAlaSerAspAsnGlyProIleThrGlnLysGlu 1603
Db 4096 CTGCTGAGGAGACAGAAAGAGTCCGCAAAATGATTTTAAAGAAAGAAAGAAAGA 4155
QY 1604 ThrSerAla 1606
Db 4156 AGTGAATCT 4164

RESULT 5
US-09-723-262-1
: Sequence 1, Application US/09723262
: Patent No. 6379912
: GENERAL INFORMATION:
: APPLICANT: Beraud, Christophe
: APPLICANT: Sakowicz, Roman
: TITLE OF INVENTION: No. 6379912el motor proteins and methods for
: FILE REFERENCE: their use

```

```

: CURRENT APPLICATION NUMBER: US/09/723,262
: CURRENT FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: US 09/572,191
: PRIOR FILING DATE: 2000-05-17
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 4757
: TYPE: DNA
: ORGANISM: Human
: US-09-723-262-1

Alignment Scores:
Pred. No.: 1,99e-75 Length: 4757
Score: 1051.00 Matches: 417
Percent Similarity: 42.34% Conservative: 304
Best Local Similarity: 24.49% Mismatches: 548
Query Match: 7.12% Indels: 434
DB: Gaps: 60

US-09-150-867-1 (1-2954) x US-09-723-262-1 (1-4757)

QY 2 SerGluGlyAspAlaValLysValArgValArgProLeuIleGlnArgGlu--- 20
DB AATGAAGGTGATGCCATCAAGTTTGTGTGCAATTCCTCCCTGCAGAAAGATCTGGG 123
QY 21 -----GlnGlyAspGlnAlaAsnLeuGlnTrpLysAlaGlyAsnAsnThrIleSerGln 38
DB 124 TCAGCTATGAGAGACAG---AACTATGCTTATCTGTCTGTCTGCACAGATCTCCGG 180
QY 39 ValAspLysThr-----LysSerPheAsnPheAsnArgValPheAsnSerHisGlu 55
DB 181 CTGCACATCCAAACCTGAGCCCAAGACCTTCACGTTTGATCATGTTGGCAGATGGATACC 240
QY 56 SerThrSerGlnIleTrpGlnGlnIleAlaValProIleIleArgSerAlaLeuGlnGly 75
DB 241 ACTGAGGAATCTGATTCGCAACTGTGCTAAAGCATTTGGAGTCTTCGATGAGCGGT 300
QY 76 TyrAsnGlnThrIlePheAlaTrpGlnIleThrSerSerGlyLysThrTrpMetMet 95
DB 301 TATTAATGATGCATCTTGTGATGACAGACCTGCTCAGGAAGACATTTACTATGATG 360
QY 96 Gly-----ThrProAsnSerLeuGlyIleIleProGlnAlaIle 108
DB 361 GGACCATCTGATCTGATTAATTTTCTCATACCTGAGAGAGATTAATCCACGAAATTTT 420
QY 109 GlnGluValPheLysIleIleGlnGluIle-----ProAsnArgGluPhe 123
DB 421 GAATATTGTTTCTTAATTTGATCGTGAAGAAAGAGCTGAGCTGGAAGAGTTTC 480
QY 124 LeuLeuArgValSerTrpMetGluIleTyrAsnGlnIleThrValLysAspLeuGlyAsp 143
DB 481 CTTTGTAAGTTCCTTATTTGAATCTACACAGCAGATATATGATCTACTG----- 534
QY 144 AspArgArgLysLysProLeuGlnIleArgGluAspPheAsnArgAsnValTyrValAla 163
DB 535 GACTCTCATCGGCTGAGCTGACTTAAGGAGCATATCAAGAGAGGAGCTTGTGTGT 594
QY 164 AspLeuThrGlnGluIleValMetValProGlnHisValIleGlnIleTrpIleLysGly 183
DB 595 GGTCCGGTGGAGCAGGAGGTAACCTCAGCTGCTGAAACCTATATCGTGTCTCGAGGA 654
QY 184 GlnLysAsnArgHisTrpGlyGlnIleThrLysMetAsnAspHisSerSerArgSerHisThr 203
DB 655 TGAGAGAAATAGACGTGGCAGTCAACATCAATGAACAGAAATGCTTAGGTCTCATAGCC 714
QY 204 IlePheArgMetIleValGlnSerArgAspArg---AsnAspProThrAsnSerGluAsn 222
DB 715 GCTTTACAAATTAACAATAGAGTCAATGAGAAAGTAATGAGATTGCAAT----- 765
QY 223 CysAspGlyAlaValMetValSerHisLeuAsnLeuValAspLeuAlaGlySerGluArg 242
DB 766 -----ATACGACCTCCCTACTCAACCTGCTGATTTAGACAGGATCTGAAGG 813

```

```

QY 243 AlaSerGlnTrpGlyAlaGlnGluIleValArgPheLysGlnGlyCysAsnIleAsnArgSer 262
DB 814 CAAAAGATACCATGCCAGAGGAGTGAATTAAGAGACAGATTAATTAATGATCA 873
QY 263 LeuPheIleGlnGlyGlnValIleLysLysLeuSerAspGlyGlnAlaGly-----Gly 280
DB 874 TTGAGCTCCCTGGGCAAGGTATTAACAGCACTTGTGCAGCTGGGGAATGAAACAGAGA 933
QY 281 PheIleAsnTrpArgAspSerLysLysLeuThrArgIleLeuGlnAsnSerLeuGlyLys 300
DB 934 CATGTTGGCTACAGAGACTCCAAACTTATCTTACAGCGGATTCCTTGGAGGTAT 993
QY 301 AlaLysThrValIleIleCysThrIleThrProValSer-----PheAspGlnThrLeu 318
DB 994 GCCAAACACGCCATTAATTCGAAATGTTATCTGATTCAGATGCTTTGGGGAACCTTA 1053
QY 319 SerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGlu 338
DB 1054 TCAACACTTAACCTTGTCCAAAGAGCCAAAGCTGATTAACAAAGCAGTAGTAAATGA 1113
QY 339 ValLeuAspArgGlnAlaLeuLeuLysArgTrpArgLysGlnIleLeuAspLeuLysLys 358
DB 1114 -----GACACCCAAAGAAATGTGAGCCAGCTCCAAAGCTGAAGTGAAGAGCTCAAGAA 1167
QY 359 GlnLeuGluAsnLeuGlnLysSerSerSerGlnTrpLysAlaGlnAlaMetAlaLysGlnGlu 378
DB 1168 CACTGGCGGAGCTTGTCTCAGAGACAGACACCAGAAACCTTCCTGACAGAGACAA 1227
QY 379 HisThrGlnLeuLeuAlaGlnIleLysGlnLeuHisLysGlnArgLysArgIleTrp 398
DB 1228 AAGAAGACTACTATATGAGAGTATTTCCAG-----GAAGCAATGTTATTC 1272
QY 399 HisLeuThrAsnIleValAlaIleSerSerGlnGlnSerGlnGlnAspGlnArgValLys 418
DB 1273 TTT-----AAGAAATCTGAACAGAAAGAAAG----- 1299
QY 419 ArgLysArgArgValThrTrpAlaProGlnLysIleGlnAsnSerLeuHisAlaSerGly 438
DB 1299 ----- 1299
QY 439 ValSerAspPheAspMetLeuSerArgLeuProGlnLysPheSerLysLysAlaLysPhe 458
DB 1300 ---TCTGTGATAGAAAGATCCCAATTAAGACAGCTCACCTCAAAAAGAAATAATT 1356
QY 459 SerAspMetProSerPheProGlnIleAspAspSerValCysThrGlnPheSerAspPhe 478
DB 1356 ----- 1356
QY 479 AspAspAlaLeuSerMetMetAspSerAsnGlyIleAspAlaGlnTrpAsnLeuAlaSer 498
DB 1357 -----ATTCATCTTAATAAATGATGTGAATTCGAA-----GAGAGT 1395
QY 499 LysValThrHisArgGluLysThrSerLeuHisGlnSerMetIleAspPheGlyGlnIle 518
DB 1396 CAAATTAATACCTTGGAAGAG-----CTCCAC----- 1422
QY 519 SerAspSerValGlnPheHisAspSerSerLysGlnAsnGlnLeuGlnTrpLysProLys 538
DB 1423 -----AAGAAATCCCGGAGAGCTTTCTCCTGAG 1452
QY 539 AspSerGlyAspMetAlaGlnCysArgLysAlaSerPheGlnLysGlnIleThrSerLeu 558
DB 1453 GAGCAGAGATCTTGTGC-----TCAAGATTAAGAAATGAGATTCATCAACTCTG 1500
QY 559 GlnGlnGlnLeuGlnSerLysGlnGlnGlnLysLysGlnLeuValGlnSerPheGlnLeu 578
DB 1501 CGAGAACAAATAGAGACACACCAGAGTTGCCAAAGTATGCTATGAAATAATCATCTCCTC 1560
QY 579 LysIleAlaGlnLeuGlnGlnGlnLeuSerValLysAlaLysAsnLeuGlnMetValThr 598
DB 1561 AGG-----GAGAGAAATAGAAAGACTGAGATTA-----TTAAGACCTGTGAAA 1602

```

QY 599 AsnSerArgLuhHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValVal 618
 : : : : :
 Db 1603 AGAGCTCAAGAA----- 1614
 QY 619 ArgLysGluMetSerValLeuGluLysPserGlyTyrAsnAlaSerAsnSerAspLeuGln 638
 : : : : :
 Db 1614 ----- 1614
 QY 639 AspSerSerValAspLysArgLeuSerSerHisAspLysIleGluHisArg 658
 : : : : :
 Db 1614 ----- 1614
 QY 659 LysMetLeuGluGlnLysIleValAspLeuGluGlnPheIleGluAsnLeuAsnLysLys 678
 : : : : :
 Db 1615 ---ATGATGCCCAAGACCTGCAAAAGCTTCTGTAATTAAGTGCATG 1671
 QY 679 SerGluAsnAspLysGlnLysSerSerGluGlnAspPheMetGluSerIleGlnLeuLys 698
 : : : : :
 Db 1672 GAGAAAGATGACAAA-----ATCAGCAAGATTTTCACTTAAGCTCAGAAAGAG 1722
 QY 699 GluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeuMetArgAspAsn 718
 : : : : :
 Db 1723 CCATGTTGTGTGCA----- 1737
 QY 719 PheAspAsnIleIleLeuGluAsnGluThrLeuLysArgGluIleAlaAspLeuGluArg 738
 : : : : :
 Db 1738 ---AACACTGACAAGTAAAGACAACTCTGCAAAATTCAGACA 1779
 QY 739 SerLeuLysGluAsn---GlnGluThrAsnGluPheGluIleLeuGluLysGluThrGln 757
 : : : : :
 Db 1780 GAGCTGAATATTCAAAGCAAGATATGAGAAATTC-----AAGAACTTACT 1827
 QY 758 LysGluHisGluAlaGluLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn 777
 : : : : :
 Db 1828 AGGAAAGGACAGTAGAATGGAATCAGACCTTCAGCTTTGCAAAAGCG----- 1878
 QY 778 AlaGluMetTyrAsnGlnAsnLeuGluLysAspLeuGluThrLysThrLysLeuLys 797
 : : : : :
 Db 1879 ---AACCTTAATCTTGAAACCTTTTGAGACA---ACAAAAGCCTGAG 1923
 QY 798 GluGlnGluIleGlnLeuAlaGluLeuArgLys---ArgAlaAspAsnLeuGln----- 814
 : : : : :
 Db 1924 CGGCAAGAAAGTT---TCTCAGCTGAATTAATTCATCTGAACACTTAAGATTATA 1977
 QY 815 ---LysLysValArgAsnPhe-----AspLeuSerValSer 825
 : : : : :
 Db 1978 ACTACACCAACCAAGGCTACCAACTTCATCCGACCAAGTACCAAAATTAAGCCCGAA 2037
 QY 826 MetGluAspSerGluLysLeuLysGluGluIlePheGlnLeuLysGlnSerLeuSerAsp 845
 : : : : :
 Db 2038 ATGGGAAGCTTTGGCTCTCTA-----TACACTGAGAAATTCATGATATTTAGAT 2085
 QY 846 AlaGluAlaValThrArgAspAlaGlnLysGluLysSerPheLeuArgSerGluAsnLeu 865
 : : : : :
 Db 2086 AATATATATTAATTAATGAGCCAGTTCTCTGAG----- 2118
 QY 866 GluLeuLysGluLysMetGluAspThrSerAsnTyrAsnGlnLysGluLysAlaAla 885
 : : : : :
 Db 2119 -----ATGAATGGAACAAGCTTTT 2136
 QY 886 SerLeuPheGluLysGlnLeuGluThrGluLysSerAsnTyrLysLysMetGluAlaAsp 905
 : : : : :
 Db 2137 GAGGCCATTTCTGGAAGCTTGAACAGTGCAGAGCAACAATGACTGCTCTTCAAGCAAA 2196
 QY 906 LeuGlnLysGlu-----LeuGlnSerAlaPheAsnGluIleAsnTyr 919
 : : : : :
 Db 2197 CTGATGAGAAGAGCAATAAAACCTTAAGCTTCAGCAGCATGTTGACAAACCTGGAACAT 2256
 QY 920 LeuAsnGluLeuLeuAlaGlyLysValProArgAspLeuLeuSerArgValGluLeuGln 939
 : : : : :
 Db 2257 CATCTCAACCAATG-----CAGAGAGCTTTTCTCATCAGAAAGAAAT--- 2298
 QY 940 LysLysValSerGluPheSerLysGlnLeuGluLysAlaLeuGluLysAsnAlaLeu 959

Db 2299 -----GATTGACCAAAACAGCAGAGAGAGCTTCTCTCAGACTGTAAGTCTCT 2346
 : : : : :
 QY 960 GluAsnGluVal---ThrCysLeuSerGluTyrLysPheLeuProAsnGluValGluLys 978
 : : : : :
 Db 2347 GAAAGCAGAGCTTCAGAGACTCAACATAAAATGACTTTTGAAGAGAGCTACATGAC 2406
 QY 979 LeuLysAsnGlnIleSerLysAlaSerGluGluIleMetLeuLeuLysGlnGlu----- 996
 : : : : :
 Db 2407 CTGCGAGTAGTCTTCATCTGCTGACAGAGAGCTTCTTCAAGTAAATTTGGAATATAGT 2466
 QY 997 -----GlyGluHisSerAlaSerIleIleSerLysGlnGluIleMet 1011
 : : : : :
 Db 2467 TCATTCAAACGAATCAGAGAGAGAGAAATTCAAACAACCTTCGAAAGACATGATGTA 2526
 QY 1012 GlnGluGlnSerGluGlnIleLeuGlnLeuThrAspGluValThrHisThrGlnSerLys 1031
 : : : : :
 Db 2527 CAGCTTCATTTAGAAATTCAGGTTAGAAAAGAAAAGCTCTTGAGAGCAAAAGCCTGC 2586
 QY 1032 ValGlnGlnThrGluGluGlnTyrLeuGluMetLysLysMetHisAspAspLeuPheGlu 1051
 : : : : :
 Db 2587 CTACAGAGATTCCTATGACAACTTACAGAAATTAATGAATTTGAGATTGACCAACTTCA 2646
 QY 1052 LysTyrIleArgAsn---LysSerGluAlaGluAspLeuLeuArgGluMetGluAsnLeu 1070
 : : : : :
 Db 2647 AGAAACCTCCAAACCTTCAAAAGAAAAGAAATGAAACTGTAATCTGATCTGAATATTTG 2706
 QY 1071 LysGlyThrMetGluSerValGluValLys-----IleAla 1082
 : : : : :
 Db 2707 ATGGACCTCTTGAGGACAAAAGAACCAATTAACAAATTTATCATTCAGTTGAGACA 2766
 QY 1083 AspThrLysHisGluLeuGluGluThrIleArgAspLysGluGlnLeuLeuHisLys 1102
 : : : : :
 Db 2767 GATTAAGAAAGACGTTCTTAAGAAATCTTAAGAAATCTTGAGCGTGTGCTGAGAGAAA 2826
 QY 1103 ---LysTyrPhePheGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSer 1121
 : : : : :
 Db 2827 CAGAAAGACAGCGCCAAAGTGTGAGCGAGCATGGCAAAAGTACAGAAACTTAAGAGAGAC 2886
 QY 1122 LeuProPheSerLysLeuValGluLysAsnSerGlnAspProIleGluIleAsnAspTyr 1141
 : : : : :
 Db 2887 TTG----- 2889
 QY 1142 HisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGlnThrGlu 1161
 : : : : :
 Db 2890 -----CTTCTACTGAAAA---GTGATGAGTTCCTCGAAAAAGTCT 2928
 QY 1162 ArgAsnSerLeuLysGluGlnValIleAspLeuAsnThrGlnLeuGlnSerLeuGlnAla 1181
 : : : : :
 Db 2929 AGAGATTTCATTAAGAAAGTGTAGCTGACCTCATGACAGATCCAGAGCTAAGATCA 2988
 QY 1182 GlnSerIleGluLysSerAsp---LeuGlnLysProLysGlnAspLeuGluGluGlu 1200
 : : : : :
 Db 2989 TCGCTGTGTGAGAAACAGAAACTATAGACACCCCTGAACAGAACTGAGAGCAATTAAT 3048
 QY 1201 ValLys-----LeuLeuGluGluMetGluLeuLeuLysGlyHisLeuThrAspSer 1217
 : : : : :
 Db 3049 TGCAAATACACTCTGCTTGTGTTGACAGAGAA-----GAGAGC 3087
 QY 1218 GlnLeuSerIleGluLysLeuGlnLeuGluAsnLeuGluValThrGluLysLeuGlnThr 1237
 : : : : :
 Db 3088 AGAGTGTGTATCAAGAAAGCAGAAAGTGAATATCTGATCTGAGAAACCCCTTAGG--- 3144
 QY 1238 LeuGlnGluGluMetLysAsnIleThrIleGluArgAsnGluLeuGlnThrAsnPheGlu 1257
 : : : : :
 Db 3145 -----CTGAGAAATCTTTTGTGAG 3162
 QY 1258 AspLeuLysAlaGluHisAspSerLeuLysGlnAspLeuSerGluAsnIleGluGln--- 1276
 : : : : :
 Db 3163 GACATA-----GAGAGGATATGCTCTGTGAGAGACTGCTCATGCCACTGAGCAGCTG 3216
 QY 1277 -----SerIleGluThrGlnAspGluLeuArgAlaLeuGlnGlu 1290

Db 3217 AACATGCTCACAGAGGCGCTCAAAAAACATCGGGCGCTGCTGAGTCTGCCAGAGAA 3276
 QY LeuArgGluGlnGlnLeuValAspSerPheArgGlnGlnLeuAspCysSerVal 1310
 Db 3277 CTGACCAAGAAAGAACCCCTATTCAGAACTTCAGCAAGCTA----- 3321
 QY GlyLeuSerProAsnHisAspAlaValAlaAsnGlnGlnValSerLeuGlyGlu 1330
 Db 3322 -----AACCAAAAGAA----- 3333
 QY ValAsnSerLeuGlnSerGluMetLeuArgGluValArgAspGluLeuGlnThrSerCys 1350
 Db 3334 -----GAGGAGGTGACACAGAGAAAGAAATGAAATATTAATCTCAAAATG 3375
 QY LysAlaLeuValSerGluLeuGlnLeuValArgAlaHisValCysSerValGluGlyGlu 1370
 Db 3376 AGGCAACTA-----GAACTGTGATGATGATTCGCTGCTGAG 3411
 QY Asn-----LeuGluIleThrLysLysLeuAsn 1379
 Db 3412 GATCCCGAGATCCTAAGACACACCTCACTTCAACACATTTGGCAAAATCCTCGAA 3471
 QY GlyLeuGlnLysGluIleLeuGlyLysSerGluLeuSerGluValLeuLysSerMetLeu 1399
 Db 3472 ACACAAAGACAGAGATA-----GAGATGGAGAGCGCTTAAGACTTCTTGTG 3519
 QY GluAsnLeuLysGluAspAsnLysLeuLysGlu-----GlnAlaGluGluTyrSer 1417
 Db 3520 GAACACCTTGTAAACAAAGCTTAATGAACAGACAGAAAGTCAAAATGCTGAAATCCTCGA 3579
 QY SerLysGlnGlnGlnPheSerLeuGluValPheSerGlySerGlyLysLeuValAsp 1437
 Db 3580 ATGAAGAGCAGATGCTGCTGAATATGAAACCTACCGCTGAAAGTCAAGCACTTAATGAG 3639
 QY GluIleGluValLeuLysAlaGlnLeuLysAlaAlaGluValGluGlnIleLysAsp 1457
 Db 3640 AAAAATCGGCTCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3699
 QY ArgAspTyrPheGluLeuValGlnThrAlaAsn----- 1468
 Db 3700 CAGAAATCATCCAGATATCAACACAGTCAAGATGAACAGAAAGATCAAAAGAAAGA 3739
 QY 1469 -----ThrAsnLeuValGluGlyLysLeuGluThrProLeuGlnAlaAspHisGlu 1485
 Db 3760 CTTCGCAAAAGTAAATAGTTGAAGAAATGCTGAAA-----ATGAAGCAGACCTGAAA 3813
 QY GluAspSerIleAspArgAspSerGluGluMetGluIleLysValLeuGlyGluLysLeu 1505
 Db 3814 GAAGTCCAAAGTCCCTTTACAAACAAAGAGATGGAATGCTTAGAATGACTGATGAAGTC 3873
 QY GluArgAsnGlnTyrLeuLeuGlnLysLeuArgLeuGlnGlnGlnLysLeuSerAsnLys 1525
 Db 3874 GAACGAACCCAAACTTTGGAGCTCAAAAGCATTCACAGAAAAAGAAAGCACTGATCAAG 3933
 QY 1526 LeuGluIleLeuGlnLysGluMetGluThrSerValLeuLysAspAspLeuGlnGln 1545
 Db 3934 CTGGAAGAAATGTATGAAGAAAGAGAGAAATCCAGAGATGGAATGTTAAGAGAG 3993
 QY LysLeuGlnLysLeuLeuSerGluAsn-----IleIleLeuLysGluAsnIleAspThr 1563
 Db 3994 CAGGTGAGTGTCTGTGAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4047
 QY ThrLeuLysHisHisSerAspThrGlnAlaGlnLeuGlnLysThrGlnGlnGlnLeuGln 1583
 Db 4048 -----CATCGAAGATTCAGTACGTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 4095
 QY 1584 LeuAlaLysAsnLeuAlaIleAlaAlaSerAspAsnCysProIleThrGlnGlnGln 1603
 Db 4096 CTTCGTAGGAGACAGAAAGTGTGCGGAAATATATATTTTAAAGAAAGAAAGAAAGA 4155
 QY 1604 ThrSerAla 1606
 Db 4156 AGTGAATCT 4164

RESULT 6
 US-09-723-219-1
 : Sequence 1, Application US/09723219
 : Patent No. 6391613
 : GENERAL INFORMATION:
 : APPLICANT: Beraud, Christophe
 : APPLICANT: Sakowicz, Roman
 : APPLICANT: Mood, Kenneth
 : TITLE OF INVENTION: No. 6391613el motor proteins and methods for
 : TITLE OF INVENTION: their use
 : FILE REFERENCE: 1017
 : CURRENT APPLICATION NUMBER: US/09/723, 219
 : PRIOR FILING DATE: 2000-11-27
 : PRIOR APPLICATION NUMBER: US 09/572,191
 : PRIOR FILING DATE: 2000-05-17
 : NUMBER OF SEQ ID NOS: 6
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 4757
 : TYPE: DNA
 : ORGANISM: Human
 US-09-723-219-1
 Alignment Scores:
 Pred. No.: 1,99e-75 Length: 4757
 Score: 1051.00 Matches: 417
 Percent Similarity: 42.34% Conservative: 304
 Best Local Similarity: 24.49% Mismatches: 548
 Query Match: 7.12% Indels: 434
 DB: Gaps: 60
 US-09-150-867-1 (1-2954) x US-09-723-219-1 (1-4757)
 QY 2 SerGluGlyAspAlaValLysValCysValArgProLeuIleGlnArgGlu--- 20
 Db 64 AATGAGGTGATGGCAATCAAAAGTTTGTGCGAATTCCTCTCGCAAGAAAGATCTGGG 123
 QY 21 -----GlnGlyAspGlnAlaAsnLeuGlnIlePheLysAlaGluAsnAsnThrIleSerGln 38
 Db 124 TCAGCTGATGAGAGACAG---AACTTATGCTTATCTGTGCTGCTCCACAGATCTGGG 180
 QY 39 ValAspGlyThr-----LysSerPheAsnPheAspArgValPheAsnSerHisGlu 55
 Db 181 CTGCACCTCAACCTCGAGCCCAAGACCTTCACGTTTGTATGATGTGGAGATGGATACC 240
 QY 56 SerThrSerGlnIleTyrGlnGlnIleAlaValProIleIleArgSerAlaLeuGlnGly 75
 Db 241 ACTCAGCAATCTGTATTCGCAACTGTGCTAAAGCAATGTGGAGATCTTCATGACGCGT 300
 QY 76 TyrAsnGlyThrIlePheAlaTyrGlnThrSerSerGlyLysThrTyrThrMetMet 95
 Db 301 TATATGTACCATTTTTCATATGATGACAGCTGCTCAGGAGAAACATTTACTATGATG 360
 QY 96 Gly-----ThrProAsnSerLeuGlnIleIleProGlnAlaIle 108
 Db 361 GGACCACTGATCTGATTAATTTTTCATTAACCTGAGAGAGATATCCCAAGATTTT 420
 QY 109 GlnGluValPheLysIleIleGlnGluIle-----ProAsnArgGluPhe 123
 Db 421 GAATATTTGTTTCTTAATGATCTGAAAGAAAGAGCTGAGCTGAAAGATTTT 480
 QY 124 LeuLeuArgValSerTyrMetGluIleTyrAsnGlnThrValLysAspLeuLeuCysAsp 143
 Db 481 CTTTGTAGTGTCTCTTATATGAAATCTACACAGCAGATATGTGATCTACTG----- 534
 QY 144 AspArgArgLysLysProLeuGluIleArgGluAspPheAsnArgAsnValTyrValAla 163
 Db 535 GACTCTGATCGGCTGGAGCTGTACTTAAGGAGCATATCAAGAGGAGAGCTTGTGTGT 594
 QY 164 AspLeuThrGluGluLeuValMetValProGluHisValIleGlnIlePheLysGly 183
 Db 595 GGTGCGGTGAGCAGGTGTGAACCTCAGCTGCAAGCTATACAGGTGTGTGTGAGGA 654

QY	184	GIuysasnaRghsTYrGlyGluThrLysMetCaspHisSerSerArgSerHisThr	203
		:::	
Db	655	TGGAGGAAATACACTGTGGCATCAACATGCATGACAGAAATCGTCTAGCTCATGCC	714
QY	204	ILepheArgMetILEValGlnSerArgAspArg---AsnAspProHisAsnSerGluAsn	222
		:::	
Db	715	GCTTTACAAATTACAAATGAGCTCAATGAGCAAAAGCTAATGAGATTGTGAAT	765
QY	223	CysAspGlyAlaValMetValSerHisLeuAsnLeuValAspLeuAlaGlySerGluArg	242
		:::	
Db	766	-----ATACGGACCTCCCTACTCAACTGGTGTGATTAGCAGGATCTGAAAG	813
QY	243	AlaSerGlnThrGlyAlaGluGlyValArgLeuGluGlyCysAsnILEAsnArgSer	262
Db	814	CAAAAAGATATCCATGCGAAAGGAGTGAATGAAGAACAGAGTAACATAAATCCGTCA	873
QY	263	LeupheILELeuGlyGlnValILELysLysLeuSerAspGlyAlaGly-----Gly	280
Db	874	TTGACCTGCCTGGGGCAAGTAGTATTACAGCACTTGTCCGCGGGGTATGTGAAGAACAGA	933
QY	281	PheILEAsnTYrArgAspSerLysLeuThrArgILELeuGlnAsnSerILEuGlyGlyAsn	300
		:::	
Db	934	CATGTTCTCTACAGAGACTCCAAACTTAACCTTCTACTACGGGATGCCCTGGAGGTAA	993
QY	301	AlaLysThrValILEILECysThrILEThrProValSer-----PheAspGluThrLeu	318
Db	994	GCCAAACAGGCATATATGCAAAATGTTATCTCTGCATCCAGGCTGTTTGGGGAAACCTTA	1050
QY	319	SerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGlu	338
Db	1054	TCAAACACTTAACCTTGTCTCAAAAGAGCCAAAGCTGATTAAAAAAGAGCAGTAGTAATGAA	1111
QY	339	ValLeuAspAspGluAlaLeuLeuLysArgTYrArgLysGluILELeuAspLeuLys	358
Db	1114	-----GACACCCAGAGAAATGTGAGCCAGCTCCACGTGAAGTGAAGAGCTCAAGAA	1166
QY	359	GlnLeuGluAsnLeuGlnSerSerSerGluThrLysAlaGlnAlaMetAlaLysGluGln	378
Db	1168	CAACGCGGAGACTGTGCTCGAGCAGACACACACAGAAAGCTTCTGCACAGAGACAAA	1222
QY	379	HisThrGlnLeuLeuAlaGlnILELysGlnLeuHisLysGluArgGluAspArgILETrp	398
Db	1228	AAGAAAGACTAACTATATGAGATATTCCAG-----GAGCAAAATTATTC	1277
QY	399	HisLeuThrAsnILEValAlaLaserSerGlnGlnSerGlnGlnAspGlnArgValLys	418
Db	1273	TTT-----AAGAAATCTCGAACGAGAAAGAG-----	1299
QY	419	ArgLysArgArgValThrTrpAlaProGluLysILEGlnAsnSerLeuHisAlaSerGly	438
Db	1299	-----	1299
QY	439	ValSerAspPheAspMetLeuSerArgLeuProGlyAsnPheSerLysAlaLysPhe	458
Db	1300	---TCTGCAATGAGAAAGATTACCCAAATTAGAAGACCTCACCCCTCAAAAAAGAAAATTT	1353
QY	459	SerAspMetProSerPheProGluILEAspAspSerValCysThrGluPheSerAspPhe	478
Db	1356	-----	1356
QY	479	AspAspAlaLeuSerMetLeuAspSerAsnGlyILEAspAlaGluTrpAsnLeuAlaSer	498
Db	1357	-----ATTCAACTATATATAATGATTGTGAATTTCCGA---GAGGAT	1399
QY	499	LysValThrHisArgGluLysThrSerLeuHisGlnSerMetILEAspPheGlyGlnILE	518
Db	1396	CAAAATAAATACGTTGGAAG-----CTCAC-----	1422
QY	519	SerAspSerValGlnPheHisAspSerSerLysGluAsnGlnLeuGlnTYrLeuProLys	538
Db	1423	-----AAGCAATCCCGGAGAGGTCTTCTGCCCTGAG	1455

QY	539	AspSerGlyAspMetAlaGluCysArgLysAlaSerPheGluLysGluIlePheSerLeu	558
Db	1453	GAGCAGGATCTCTTGGTC-----TCGAAATTAGGAAGAAATCAATCCATCTG	1500
QY	559	GlnGlnGlnLeuGlnSerLysGlnGlnLysGlnLeuValGlnSerPheGlnLeu	578
Db	1501	CGAGAACAAATAGAGCACCCACCCAGAGTTGCAGAGTATGCTATGAAATCAATCCCTC	1500
QY	579	LysIleAlaGluLeuGlnGlnLeuSerValLysAlaLysAsnLeuGluMetValThr	598
Db	1561	ACG-----GAGGAGAAATAGAAAGCTGAGATT-----TTAGAGCGCTGTAA	1602
QY	599	AsnSerArgLysIleSerIleAsnAlaGluValGlnThrAspValGluLysGluValVal	618
Db	1603	AGAGCTCAAGAA-----	1614
QY	619	ArgLysGluMetSerValLeuGlnAspSerGlyTyrAsnAlaSerAsnSerAspLeuGln	638
Db	1614	-----	1614
QY	639	AspSerSerValAspGlyLysArgLeuSerSerHisAspGluLysIleGluHisArg	658
Db	1614	-----	1614
QY	659	LysMetLeuGlnGlnLysIleValAspLeuGlnGlnPheIleGluAsnLeuAsnLysLys	678
Db	1615	--ARGATGCCCCAGACCATCTGCACAACTAGAAAAAGCTTCTCGAATAAGATGCGATG	1671
QY	679	SerGluAsnAspLysGlnLysSerSerGlnGlnAspPheMetGluSerIleGlnLeuLys	698
Db	1672	GAGAAAGATGCAAA-----AATCAGCAAGAGATTTTCACCTTAAGCTCAGAAAGAG	1722
QY	699	GluAlaIleMetAlaGluLysAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAspAsn	718
Db	1723	CCATGTTTGTTGCA-----	1737
QY	719	PheAspAsnIleIleLeuGluAsnGlnIleThrLeuLysArgGluIleAlaAspLeuLysArg	738
Db	1738	-----AACCTGAGAAAGTTAAAGCAACCTCTCGAATAATGACGA	1779
QY	739	SerLeuLysGluAsn-----GlnGluThrAsnGluPheGluIleGluGluLysGluThrGln	757
Db	1780	GAGCTGAATTAATTCAAAGCAAGAAATAGAGAAATTC-----AAAGAACTTACT	1827
QY	758	LysGlnHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn	777
Db	1828	AGGAAAGCAGAGCTAGAAATGGAATGCAGACTTCAGCTCTTCCAAAAACGC-----	1878
QY	778	AlaGluMetTyrAsnGlnAsnLeuGlnLysAspLeuGlnIleThrLysThrLysLeuLys	797
Db	1879	-----AACCTTAATCTTGAAACCTTTTGGAAGCA--ACAAAGCCTTGCAAG	1933
QY	798	GlnGlnGlnIleGlnLeuAlaGluLeuArgLys--ArgAlaAspAsnLeuGln-----	814
Db	1924	CGGCAAGAAATT-----TTCAGCTCAATTAATTAATTCATGCTGAAACCTTAAGTTATA	1977
QY	815	-----LysLysValArgAsnPhe-----AspLeuSerValSer	825
Db	1978	ACTACACCAACCAAGCGCTACCAACTTCATTCGCCAGACAGTACCAAAATTAAGCCCTGAA	2037
QY	826	MetGlyAspSerGluLysLysCysGlnGluIlePheGlnLeuLysGlnSerLeuSerAsp	845
Db	2038	ATGGAGAACTTTGGCTGCTCTA-----TACACTCAGAAATTCGTGCAATTATTAGAT	2085
QY	846	AlaGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerLysAsnLeu	865
Db	2086	AATGATATATTAAATGAGCCAGTTCTCTGAG-----	2118
QY	866	GluLeuLysGluLysMetGluAspThrSerAsnTyrAsnGlnLysGlnLysAlaAla	885
Db	2119	-----ATGATATCAACAACCTTTT	2136
QY	886	SerLeuPheGluLysGlnLeuGlnIleThrGluLysSerAsnTyrLysLysMetGluAlaAsp	905


```

QY 1564 ThrLeuLysHisSerAspThrGlnAlaGlnLeuGlnLysThrGlnGlnGlnLeuGln 1583
DB 4048 -----CATCGAAGATTCAGTACGTAGTGGCATTAAGAGAAATGTCAGG 4095
QY 1584 LeuAlaLysAsnLeuAlaIleAlaIleSerAspAsnGlyProIleThrGlnGlnLysGln 1603
DB 4096 CTGCTGACGAGACAGACAAAGTTGCTGCCGAAATGTATTTTAAAGAAAGAAAGAGA 4155
QY 1604 ThrSerAla 1606
DB 4156 AGTGAATCT 4164

RESULT 7
US-09-592-054-7
; Sequence 7, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finet, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684el motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4127
; TYPE: DNA
; ORGANISM: Human
US-09-592-054-7

Alignment Scores:
Pred. No.: 1.38e-66 Length: 4127
Score: 941.00 Matches: 357
Percent Similarity: 42.37% Conservative: 243
Best Local Similarity: 25.21% Mismatches: 449
Query Match: 6.37% Indels: 368
DB: 4 Gaps: 53

US-09-150-867-1 (1-2954) x US-09-592-054-7 (1-4127)
QY 3 GlnGlyAspAlaValLysValCysValArgValArgProLeuIleGlnArgGlu----- 20
DB 118 AAGGGAATTCCTGTAAGAGTGGCAGCTGCTTGCCTGTGTCGCCCAAGAGATTAGC 177
QY 21 GlnGlyAspGlnAlaAsnLeuGlnTrpLysAlaGlnLysAsnThrIleSerGlnValAsp 40
DB 178 GAGGGGCGCCAGATGTCCTTCCTCCGCCGCGGAGAGACTCAGTGG-----GTGCTT 231
QY 41 GlyThr-----LysSerPheAsnPheAspArgValPheAsnSerHisGlnSerGln 59
DB 232 GGTTCATATTAAGCAGTAGCGCGCTCATTAAGGCGCATATTAAAGATTAATGCAAG 351
QY 60 IleTyrGlnGlnIleAlaValProIleIleArgSerAlaLeuGlnGlnIleTyrAsnGlyThr 79
DB 292 GTCTTCATATTAAGCAGTAGCGCGCTCATTAAGGCGCATATTAAAGATTAATGCAAG 351
QY 80 IlePheAlaTyrGlnIleThrSerSerGlyLysThrTyrThrMetGly----- 96
DB 352 GTCTGCGCTATGGGCGAGAGTGGCTCGGAAAAACCTATTCATATGAGGAGTGCATACAT 411
QY 97 -----ThrProAsnSerLeuGlyIleIleProGlnAlaIleGlnGlnValAlaPhe 112
DB 412 GCGGACAGAGAGATGAGACCAACAGTTGGCATTAATTCAGGTAATACAGTGCCTTTC 471
QY 113 LysIleIleGlnIleLysProAsnArgLysPheLeuLeuArgValSerTyrMetGlnIle 132
DB 472 AAGGAATATGATCAAAAGAGTGAATTTGATTTACTCTGAAAGTCTTCTTACTAGAGATT 531

```

```

QY 133 TyrAsnGlnThrValLysAspLeuLysCysAspAspArgArgLysLysProLeuGlnIle 152
DB 532 TACATGAGAAAGAAATTTTGGATCTTCTATAGCCATCTCGTGAGAAAGCTCAAAATTAATTA 591
QY 153 ArgGlnAspPheAsnArgAsnValTyrValAlaAspLeuThrGlnGlnLeuValMetVal 172
DB 592 CCGGAGAGATCCTTAAGGAGAGCATTAAGATGTGGGACTCTCAGAGAGACTGTGTTAGTT 651
QY 173 ProGlnHisValIleGlnTrpIleLysGlyGlnLysAsnArgHisTyrGlnGlnIle 192
DB 652 GCCTTGGATCTCTTTCCTGTTGGACGAGGACACACTCTAGACTGGGCTCCACAC 711
QY 193 LysMetAsnAspHisSerSerArgSerHisThrIlePheArgMetIleValGlnSerArg 212
DB 712 GGTATGAACTCCAGTCGCTCCGATCCATGCATCTTTCATATCTTCCTAGAGCAAGGA 771
QY 213 AspArgAsnAspProThrAsnSerGlnAsnGlyAsnGlyAlaValMetValSerHisLeu 232
DB 772 AAGAAAAGTACAAAGATACAGC-----TTTCGCTCCAAAGCTG 810
QY 233 AsnLeuValAspLeuAlaGlnSerGlnArgAlaSerGlnThrGlyAlaGlnLysValArg 252
DB 811 CATCTTGTACACTCGCTGATCAGAAAGACAGAAAGAAAGCAAGCTGAGAGGATCGT 870
QY 253 LeuLysGlnGlyCysAsnIleAsnArgSerLeuPheIleLeuGlnGlnValIleLysLys 272
DB 871 CTAAAGAGGGTATTATATTATTAACCGAGCCCTCTGATGCTGGGAATGTAATCACTGCT 930
QY 273 LeuSerAspGlyGlnAlaGlyGlyPheIleAsnTyrArgSerLysLeuThrArgIle 292
DB 931 CTGGAGATGACAAAAGGCTAGCTTGTGCTCCTACAGAGATTCAGAACTGACAGCTG 990
QY 293 LeuGlnAsnSerLeuGlnGlyGlnAlaValThrValIleLeuCysThrIleThrProVal 312
DB 991 CTGCAAGATCTCTAGAGGTAAAGCAGCACCTTATATAGTACCGTGTGATCCTGCT 1050
QY 313 -----SerPheAspGlnThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHisVal 330
DB 1051 GACTCCAAATCTAAGGAAACATTAAGTACCTTCCTGATCTGACAGAGCAAGAAATAAT 1110
QY 331 ArgAsnThrProHisValAlaGlnGlnValLeuAspAspGlnAlaLeuLysArgTyrArg 350
DB 1111 AAGAACAAACCTATTGTGTAAT-----ATTGATCCCCACACA 1146
QY 351 LysGlnIleLeuAspLeuLysLysGlnLeuGlnAsnLeuGlnSerSerSerGlnThrLys 370
DB 1147 GCTGAAGTAAATCATTAAGAAAGCAAGGTACACAGCTA----- 1185
QY 371 AlaGlnAlaMetAlaLysGlnGlnHisThrGlnLeuLeuAlaGlnIleLysGlnLeuHis 390
DB 1186 -----CAAGCTCTGTTG----- 1197
QY 391 LysGlnArgLysPargIleTrpHisLeuThrAsnIleValAlaSerSerGlnGlu 410
DB 1197 ----- 1197
QY 411 SerGlnGlnAspGlnArgValLysArgLysArgArgValThrTrpAlaProGlyLysIle 430
DB 1197 ----- 1197
QY 431 GlnAsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeuProGly 450
DB 1198 -----CTACAAACCCATGGA-----GTTACCCGCTCTGGA 1227
QY 451 AsnPheSerLysLysAlaLysPheSerAspMetProSerPheProGlnIleAspAspSer 470
DB 1228 TCTATTAATGACAGAA-----CCATCA----- 1248
QY 471 ValCysThrGlnPheSerAspPheAspAlaLeuSerMetMetAspSerAsnGlyIle 490
DB 1249 -----GAGAACTCTCAATCCCTGATGAGAGAAAT----- 1278
QY 491 AspAlaGlnIleTrpAsnLeuAlaSerLysValThrHisArgGlnLysThrSerLeuHisGln 510

```


Db 1279 -----CAG 1281
 QY 511 Sermettleaspheglycylleaseraservalglnphehisaspserserlysglu 530
 Db 1282 TCCCTGGTAGAG-----GAGAAATGAA 1302
 QY 531 Asnglnleuglnlyrleuprollysaspserglyasmetalaglucylargylsalaaser 550
 Db 1303 AAATTAACTGTGTCTGAGCAAGCGAGCTGTGACAGACCCAGATGTTGGAGAGGATC 1362
 QY 551 Phegllysgluilethrserleuglnleuglnleuglnserlysgluugluyllys 570
 Db 1363 ATTTGACAGAG-----CAAGTGAAATGAAACAGCAAGCCAGCACTGAGAGAGCTCAGG 1416
 QY 571 Glu-----leuvalglnserphegluleulysile 580
 Db 1417 CAGCATGCGCGCTGCACACCTGATCTCAAAAGCTAGTGAGACTTGTGAGAGACCG--- 1473
 QY 581 AlaagluleuglnleuglnleuSERVallysAllysAsnleuglnmetValthranser 600
 Db 1474 ---GATTTGAAAGAAATGTAGAGATATTGTAACTGCAGCACTGATTAACCACTTA 1530
 QY 601 ArggluHisserile-----AsnAlaglulvalglnthrAspvalglulysgluval 617
 Db 1531 TCAGATGAAACTGTGCTTGACAGCGCTGCACCATTTGACTGCGGTGAGAGAGAA--- 1587
 QY 618 ValArgylsleumetSERValleuglnlyaspserglytyrAsnAlaserAsnserAspleu 637
 Db 1588 -----GCTCAAGTGGAAACCACT 1605
 QY 638 GlnaspserservalAspglyllyargleuSERSerHisAspglylucylleuglnHis 657
 Db 1606 CCAAGACACACAGAGTCTTCTGACGCTTTTACACTCAGCAT-----GCTCTCAT 1656
 QY 658 ArgylsleumetleuglnlylvalAspleuglnleupheile-----glu 673
 Db 1657 CAACCTAGATGCTGAAGAGAGGTGTGATGTAATAAGCCCTTGCACTCAAAAGAGCC 1716
 QY 674 AsnleuAsnlyllysserGlnasAspAspAspAspAspAspAspAspAspAspAspAsp 693
 Db 1717 CTACTTGGAAAGATGACTCAAGACACCAACCA-----CTACAG 1755
 QY 694 SerlleuglnleucysgluAlalleuAlaaglulysAlaAsnAlaAlaAlaAlaAla 713
 Db 1756 CCCATTCAGTT----- 1767
 QY 714 LeuMetArgasAsnAspAsnIlelleuglnleuglnleuglnleuglnleuglnleuglnle 733
 Db 1768 CAATACACAGATTAATTAATAATCTAGAAATTAAGATCAATCAATCTCAAAAGAGAAAG 1827
 QY 734 AlaAspleuglnlyargserleu-----lysgluAsnleuglnleuglnleuglnleuglnle 750
 Db 1828 GAAATATTTGTTCTGAACTTCAAGACAGCAAGAAAGAAATGCCAACCAAGCCAGCTGAGT 1887
 QY 751 lleuglnlysglnthrGlnlysglnHisgluAlaAlaAlaAlaAlaAlaAlaAlaAla 764
 Db 1888 GAGCAACGCTGCCAAACTTCTCAGAGAGCTGAGAGGTCAAAATAGCTGATCAAGAGAA 1947
 QY 765 lleHisgluileglyserleuLysleuValgluAsnAlaAlaAlaAlaAlaAlaAlaAla 784
 Db 1948 CTGAATAGAGCTCAAACTTCTGAACCTAAAGATCCACAGAGAGCTGCTCTCCAAA 2007
 QY 785 leuglnleuglnleuglnleuglnleuglnleuglnleuglnleuglnleuglnleuglnle 804
 Db 2008 CTGACACAGAGGATA-----CGATGTGAAAGAAACCGCGGTGACAGTTA--- 2052
 QY 805 GluLeuArgylsArgAlaAspAsnleuglnlyllyllyllyllyllyllyllyllyllyllyl 824
 Db 2053 ---ATGGGTCAAAATGAAGAGAGATGCTGAG---AAGTTTACACAGTGS----- 2094
 QY 825 Sermetgllyaspserglyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 844

Db 2095 -----AACGAGAAAGACAAAGAAATACAGTTAAAGAA----- 2133
 QY 845 AspAlaAlaValThrArgasPalaglulysglucylserpheleuArgsergluAsn 864
 Db 2134 -----CGAGACCGTAAAGAGCA-----TATACGCTG 2160
 QY 865 leuglnleuglnlyllysmetgluAspThrserAsnTPtyrAsnleuglnlyllyllyllyl 884
 Db 2161 CTGAATCTTGAAGAAACTTCCAGAAACATCAATGCTGCTCAGCAAGTAAACGAGAG 2220
 QY 885 AlaSerleupheglulysgluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 904
 Db 2221 GCAGCAGCTGCCAACACAGCGTCTCAAGATGCTCTCCAGAAACACGAGAGGTT---GCA 2277
 QY 905 Aspleuglnlysglnleuglnleuglnleuglnleuglnleuglnleuglnleuglnleugln 923
 Db 2278 GATAAGCGGAAAGACACACAGAGC-----CGTGAAATGGA 2313
 QY 924 ---leuAlaGlyLysValProArgAspLeuSerArgValgluLeuglnlyllyllyllyl 942
 Db 2314 GGCACCTGCAGCTCGAGTAGAGAAATGGCTTGAATGAATGAATGAGTTATGCTGACT 2373
 QY 943 SergluPheSerLysgluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 962
 Db 2374 GAGGAAGCCAAACGACATCTGAATGACCTCTTGAAGACAGAAAGATGCTGCTGAGAT 2433
 QY 963 ValThrCysleuSergluTyrllyllyllyllyllyllyllyllyllyllyllyllyllyl 982
 Db 2434 GTGTTCAACTCAAGAA----- 2451
 QY 983 lleSerlyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 1002
 Db 2452 -----AAAAAGAAATCGGAGAAATCCACTCT 2481
 QY 1003 llelleSerlyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 1022
 Db 2482 AAACCCGGAAGTACATCTCCCTTCTGAGCGTACATGTCAGTGGATTTGGACAGAA 2541
 QY 1023 AspgluValThrHisThrGlnserlyllyllyllyllyllyllyllyllyllyllyllyllyl 1042
 Db 2542 GATTTGATTAACA-----AAACGATTTGAACCTTGAACCTGAAATGGAAGCTC 2589
 QY 1043 lyslyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 1061
 Db 2590 AGGAGCTGCAGATGCTGCAGCTACACAGAAAGCTGCTGAGTGCAGAAAGTGAATGAG 2649
 QY 1062 AspleuLeuArggluMetgluAsnleuLysgluLysleuLysleuLysleuLysleuLysleu 1081
 Db 2650 CG-BCAAACAAATGCTGGAGAAATATTCACACATCTGGAAGCC-----AAGTGT 2699
 QY 1082 AlaAspThrlyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 1101
 Db 2700 GCC---CTGAATATTTGATTTGAGAGAGCTGCTCTCTCCAAA-----ATACAT--- 2744
 QY 1102 lyslyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 1121
 Db 2745 -----GTCACCAAACTGMAAAACG 2765
 QY 1122 leuProPheSerlyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 1141
 Db 2766 CTGACACAGACCAAGCGAGCTGTGCTGACATGCAAGAGATGCTA---TTTGAGAACAA 2822
 QY 1142 HisAsnleuIleAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 1161
 Db 2823 AATCATTTTCTGAGATAGACAGACAGTACAGCTGAGCTGCTGACAAATGAGACACAG 2882
 QY 1162 ArgAsnSerleuLysgluAlaValIleAspLeuAsnThrIleAspLeuSerleuLysleuLys 1180
 Db 2883 CAC-----CAAGAGAAAGGTGTATACCTTGACACAGCTGACAGAAAGCAAAATG 2933
 QY 1181 ---AlaGlnSerlleuglnlyllyllyllyllyllyllyllyllyllyllyllyllyllyllyl 1187
 Db 2934 GCAGAGAGCAGTTAGAGAAATCAGCCAGTGAAGAAAGAACCAAGTGTGTGACACACTG 2993

QY 1188 -----AspleuInlyrsProlysgInAspleuGlucInlygluVallys 1202
 DB 2994 CAGGTGCAGATGAGAACTTGAGAAAGATCGGAGATGTGGAGAAATACAG---CAG 3050
 QY 1203 LeuLeuLeuGlucInleuLeuLeuLyglYHisLeuThraspSerGlnLeuSerIleGlu 1222
 DB 3051 CTTCTCCAGAGAAATGCAATCAACAGAAACATGATCTCTCCAGCGTACAGCAGACA 3110
 QY 1223 LysLeuGlnLeuGlnAsn-----LeuGluValThrGluLysLeuGlnThrLeuGln 1239
 DB 3111 CAGAAACATCTTCTTATATATACCTTCTTATCTCCAGACTCTTCTTTAAATATATCCA 3170
 QY 1240 GluGluMetLysAsnIleThrIleGluArgAsn-----GluLeuGlnThrAsnPhgGlu 1257
 DB 3171 CCAAGACCAAAACCTCTGCTGTAAAGAAAGTTCTGGAGCAAGATGACATCGACAG 3230
 QY 1258 AspleuLys-----AlaGlnHisAspSerLeuLysGlnAspleuSerGlnIleGlu 1275
 DB 3231 GATCTAAATATATGTCTCAAGATCTCTGTGAATGACCAAGATGATGTGATGCG 3290
 QY 1276 GlnSerIleGluThrGlnAspGlu-----LeuArgAlaIleGln 1288
 DB 3291 GACATGATGAGGGGATGATGAGCAATGAGAACCCAAATAATTAATCTCAAG 3350
 QY 1289 GluGluLeuArg-----GluGlnLysGlnLeuValAspSer 1300
 DB 3351 AAGAACATCCAAAGGATGCTTCGTCAGAGGGGTGTGGAGCAAGACAGTGT-----GGG 3404
 QY 1301 PheArgGlnGlnLeuLeuAspLysSerValGlyIleSer-----SerProAsnHisAspAla 1319
 DB 3405 TCGAGAACCAAAAGTCAACACTGTGTGTGACTGTGTGACCCCAACAAAGTGTGCG 3464
 QY 1320 ValAlaAsnGlnGluValSerLeuGlnValAlaSerLeuGln 1335
 DB 3465 AACCCGACCAAGCAAGATACCTTGGGCACTGTGAACAGACCAG 3512
 RESULT 8
 US-09-722-139-1
 ; Sequence 1, Application US/09722139
 ; Patent No. 6355471
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Freeman, Richard
 ; TITLE OF INVENTION: NO. 6355471el motor proteins and methods for
 ; FILE REFERENCE: 1055
 ; CURRENT APPLICATION NUMBER: US/09/722,139
 ; CURRENT FILING DATE: 2000-11-24
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 4176
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-722-139-1
 Alignment Scores:
 Pred. No.: 1.65e-64 Length: 4176
 Score: 915.50 Matches: 416
 Percent Similarity: 40.598 Conservative: 266
 Best Local Similarity: 25.518 Mismatches: 536
 Query Match: 6.208 Indels: 433
 DB: 4 Gaps: 71
 US-09-150-867-1 (1-2954) x US-09-722-139-1 (1-4176)
 QY 6 AlaValLysValcysValArgValArgProLeuIleGlnArgGluGlnGlyaspGlnAla 25
 DB 7 TCGGTCAAGGTGGCGGTGAGGGTCCGCGCCATGAATCGCAGGAAAGACTGTGAGGCC 66
 QY 26 Asn-----LeuGlnTrpLysAlaGlyAsnAspThrIleSerGlnVal----- 39

DB 67 AAGTCATTATTCAGATGGAGAAAGCAAAACGACAAATCACTTAAGATACCAGAA 126
 QY 40 -----AspGlyThrLysSerPheAsnPheAspArgValPhe 51
 DB 127 GGAGGCACTGGGAGACTCAGAGAGAGACGACCAAGACCTTACCTTACCTTCTTT 186
 QY 52 AsnSerHisGluSerThrSer-----GlnIleArgGlnIle 64
 DB 187 TATCTGTGATCAAAAGCCAGATTAAGTTTACAGAAAGATGGTTTCAAAACCTTC 246
 QY 65 AlaValProIleIleArgSerAlaLeuGlnGlyIleArgAsnGlyThrIlePheAlaTrpGly 84
 DB 247 GGCACAGATGTCTGTAAGTCTGATTTGAAGTTTAAAGCTTGTCTTTCATATGGG 306
 QY 85 GlnThrSerSerGlyLysThrThrThrMetMetGlyThrProAsnSerLeuGlyIle 104
 DB 307 CAATCGATCTGGAAGAGCATACATGATGAGGAAATTCGAGATTCGTGCTTAATA 366
 QY 105 ProGlnAlaIleGlnGluValPheLysIleIleGlnGluIle-----ProAsnArg 121
 DB 367 CCTCGATCTGGAAGGACTCTTCAGTCGATTAATGAAACCAACAGATGGATAGACT 426
 QY 122 GluPheLeuLeuArgValSerThrMetGluIleIleArgAsnGlyThrValLysAspleuLeu 141
 DB 427 TCTTTTCGACCTGAACTGACCTACTTACAAATTTATACGAACGCTGTGACATCTACT 486
 QY 142 CysAspAspArgArgLysLysPro-----LeuGluIleArgGluAspPheAsn 157
 DB 487 -----CGGGGAAGATCATCTAAACCTTCATTTGAGAGCTCGGTGATCCCAAA 537
 QY 158 ArgAsnValIleValAlaAspleuThrGluGluLeuValMetValProGlnHisValIle 177
 DB 538 GAAGGCCCTTATGTGAGATGATTCACAAATTTAGTACAGAAATTANGTACAGTAA 597
 QY 178 GlnTrpIleLysGlyGlyLysAsnArgHisLysTrpGlyGluThrLysMetAsnHis 197
 DB 598 GAACCTATGATGAGCGGAGCAATATCAACCGACCGACGACGATGGATGAACGACGTC 657
 QY 198 SerSerArgSerHisThrIlePhe-----ArgMetIleValGluSerArgAspArgAsn 215
 DB 658 AGTAGAGCTCATCCATCTTCACCAACCAAGTTCCACTCAGGCTAAATTTACTGTGA 717
 QY 216 AspProThrAsnSerGlnLysAspGlyAlaValMetValSerHisLeuAsnLeuVal 235
 DB 718 ARGCCA-----TGTGAA-----ACGTAAGTAAATTCACCTTGGTT 753
 QY 236 AspleuAlaGlySerGluArgAlaSerGlnThrGlyAlaGluGlyValArgLeuLysGlu 255
 DB 754 GATCTGCGCGGAAGTGAAGGTGAGATGCGACCGGAGCCACCGGGGTATGCTAAAGGAA 813
 QY 256 GlyCysAsnIleLeuAsnArgSerLeuPheIleLeuGlyGlnValIleLysLysLeuSerAsp 275
 DB 814 GGGGGAATATTTAAACAAGTCCCTTGACTGTGGGGAAGCTATTTCTCCCTTACGTAT 873
 QY 276 GlyGlnAlaGlyGly-----PheIleAsnTrpArg 285
 DB 874 TTATCTCAGAGTCTCAAAATACCTTTCGAAAGAACAGCAAGTTTCTGCTTACAG 933
 QY 286 AspSerLysLeuThrArgIleLeuGlnAsnSerLeuGlyLysAlaLysThrValIle 305
 DB 934 GATTCGTGTGACTGTGTGTTAAAGATAGCTTGGAGGAACCTAAACTATACAG 993
 QY 306 IleCysThrIleThrPro-----ValSerPheAspGluThrLeuSerThrLeuGlnPhe 323
 DB 994 ATTGCACACATTTACCTCTGCTGATTTATGAGAAACCTTAAGTACTTCTTCCCTAT 1053
 QY 324 AlaSerThrAlaLysHisValArgAsnThrProHisValAsnGluValLeuAspAspGlu 343
 DB 1054 GCAAAATGAGCCAAACATCATCAACAAGCCTTACCATTAATGAGATGCCAACGTCANA 1113
 QY 344 AlaLeuLeuLysArgTrpArgLysGluIleLeuAspLeuLys----- 357
 DB 1114 ----CTTATCCGTGACGTGGAGCTGAATATGCAAGACTGAAACGCTGCTGTCAAGG 1170

OY	356	LysGlnLeuGlnAsnLeuGlnSerSerGlnThrLysAlaGlnAlaMetAlaLysGlu	377
Db	1171	AAATCAAGTATGCCCTTACAGTACCCCAAC	1215
OY	378	GlnHisThrGlnLeu---LeuAlaGlnIleLysGlnLeuHisLysGlnIleArgLysArg	396
Db	1216	GAATAACTCAAGCAATGAAGCAAGGTTCAAGATTATCCAGCAATGCAAAATAG	1275
OY	397	IleThrPheLeuThrAsnIleValAlaSerSerGlnLysSerGlnAsnArg	416
Db	1276	--TCGATGAAACC-----	1293
OY	417	ValLysArgLysArgValThrThrAlaProGlyLysIleGlnAsnSerIleHisAla	436
Db	1294	ATTTGAAAGACAAACTCTA-----	1326
OY	437	SerGlyValSerAspPheAspMetLeuSerArgLeuProGlyAsnPheSerLysAla	456
Db	1337	GAAGGAGATTGGA-----	1344
OY	457	LysPheSerAspMetProSerPheProGlyIleAspAspSerValCysThr	473
Db	1345	TTGCAATCTGAAGTCCTCATTTGATTGGCATGAGTGAACCTTTTGAGTCTGGAAATC	1404
OY	474	-----GluPheSerAspPheAspAlaLeuSerMet	484
Db	1405	ATCTTATATCATTTAAAGAAAGGTCAAGACATACGTTGGTAGAGACAGATGCTTCCACGAG	1464
OY	485	MetAsp-----SerAsnGlyIleAspAlaGlu-----	496
Db	1465	CAAAATATTGTTCTTCAAGCTGACCTTGACCTTGAGAGAGATGCATCTTGAAATATTC	1524
OY	497	AlaSerLysValThrHisArgGlyLysThrSerLeuHisGlnSerMetIleAspPheGly	516
Db	1525	GGGGGACAGAGAT---CTGATACCCCTGAGTGGGTCCAACTGCTCTGTGAATGCTTT	1561
OY	517	GlnIleSerAspSerValGlnPheHisAspSerSer-----	531
Db	1582	CAGATCTGGGAGGCGCACACATCTTAATCAAGCTGCTGATTCTCTGGGAAGAACAT	1641
OY	532	GlnLeuGlnThr---LeuProLysAspSerGlyLysPheAlaGlnCysArgLys-----	548
Db	1642	ATGTTGGCTTTAAACCATCCAAAGGAAGCGCCAGACTAGGAGAAAGAGAAAGTGGC	1701
OY	549	-----AlaSerPheGlnLysGlnIleThrSerLeuGlnGlnLeuGlnSerLysGlu	566
Db	1702	CTTCTGCTCCCTTCACTCTTGCATGACCCGACCTC-----	1746
OY	567	GlnGluLysLysGlnLeuValGln-----	580
Db	1747	CGTAGAAGACTGTCGACGATGTTGTATTAACCCCGGACTTGAATTTGAAAGGCACAG	1806
OY	581	AlaGlnLeuGlnGlnGlnLeuSerValLysAlaLysAsnLeuGlnMetValThrAsnSer	600
Db	1807	CGTAGAAGACTGTAATAATTAGAAACTAAAGAAACTATAGAAATG---GAGGA	1863
OY	601	ArgGlnHisSerIleAsnAlaGluValGln---ThrAspValGlnLysGlnValValArg	619
Db	1864	AAGCAAGAAATCAGCAAGGCTGAACTGAGACGCGATGACAGAGAGAGTGAACCCACGCG	1923
OY	620	LysGlnMetSerValLeuGlnLysAspSerGlyThrAsnAlaSerAsnSerAspLeuGlnAsp	639
Db	1924	AAGGACGACAAATCGT-----	1950
OY	640	SerSerValAspGlyLysArgLeuSerSerSerHisAspGlyCysIleGlnHisArgLys	659
Db	1951	-----ATTGCCAAGCAAGAGAGAGAGAGGCTCAAAAGCGCGACG	1986
OY	660	Met---LeuGlnGlnLysLysIleValAspLeuGlnGluPheIleGlnAsnLeuAsnLysLys	678
Db	1987	TTCCACATCTCAGAAACAGCTTAAAGGATTTTA-----	2028

OY	679	serGIuaasaprysgInlysserSerGIuInlaasphemecluserIleelInleucys	698
		
Db	2029	GAAATATTGAAAGGAGAGGCTGAGGGAACAG	2061
OY	699	GIuaIlelmetlaGIuLyAlaasnaIaleuclInleuIaleumetarGapsan	718
		
Db	2062CAGGAATTCAGCTGCAGAAAGAGCA	2088
OY	719	pheaapnaIleIleuclInlaasngIuThrleuLyasrgIuIlealIaasleuIarg	738
		
Db	2089	-----CAAGAAAGAGAGACCTTCTCCGCGCTCCAGAAAGAACCTCCAGCA	2133
OY	739	-----SerleuLysgIuaasngInluThrAsngIuPhlegIuIleleuclInlys	754
		
Db	2134	CTCAAGAAAGCTCAACMACACGAGAAAGGCTGACAGAGTTTCAGATTATTTTCAAACTGGAC	2193
OY	755	GIuThrGIuLysgIuHIsGIuIaInleuIleHIsGIuIleGIserleuLyysleu	774
		
Db	2194	CAGCTCCAAAAGGAAAAAGATGAACACTATGCC---AACTTCMACTGGAAAAAAAAGCA	2255
OY	775	ValGIuaasnaIaGIu-----MetTyraasngIuaasleuclInluIaasleuGIu	790
		
Db	2251	CTAGAGGAGCAGAGAAAGAGACAGCTCATCTCTGCGCCATCTGTGAAGACAGCTCCGA	2310
OY	791	ThrLyS-----ThrLySleuIleuLyGIuInglInglIleIleuIaIaGIuIeury	807
		
Db	2311	GAGAAAGCAGAGATGATTCACACTCTCTCGCGGTGGAGAGTAAAGGTGGAGAGAGAG	2370
OY	808	LyasrGAlaIaasnaIeugInLyLyValaIargAsnPhaaspreuserValsermetGIy	827
		
Db	2371	AAGAG-----	2376
OY	828	AspSerGIuLyseuCysgIuGIuIlePhlegInleuLygsInserleuser-----	844
		
Db	2377	GACCTGGAAGGCAATTCCGGAATCCCTCTCGGAGTGAAGAGCCTGCTCCGAGGAGAT	2436
OY	845	---AspIaGIuIaIaValThrArgAspAlaInglInlyGIySerPhleuIargserGIu	863
		
Db	2437	GAAAGTGGCGAGGAGTTTGAAGAAAGCTCAACCTCGTTTCTCGAATTCAGAGAAAGGAG	2496
OY	864	AsnleuGIuIeulysGIuLyMetGIuAspThrSerAntPyraasngInlysgIuLyS	883
		
Db	2497	CTTGCAAGCTAGTGAACCTTGAGAAAGACCTGTTCCAGCAGAAAGACATCTCGAAAAA	2556
OY	884	AlaIaIaSerleuPhlegIuLysgInleuGIuThrGIuLySerAntPyraLySMetGIu	903
		
Db	2557	GAACTCCAAAGAAAGACAGAGATCTCTAGACTGTTTAAATGTGCACATCAGCAAGAAATCT	2616
OY	904	AlaAspIeugInLySgIuIeugInserAlaPhaasngIuIleAsnTyIleuAsnGIu	923
		
Db	2617	AGATTGTTGGAAAAACATGATGAGAGTCCACAGAT	2652
OY	924	LeuIaGIuLyValProArgAspIeuserArgValGIuIeugInlyLyValSer	943
		
Db	2653	---GTCCAGGAAAGTCCCTCAAGATTCGAGAAAAATAAACCCAGTGGAGTACAGGCTCGAA	2709
OY	944	GIuPheserLySgInleuGIuLySAlaIeugIuGIuLyAsnaIaIeugInuAsngIuVal	963
		
Db	2710	TATTAAGAACCCAGCTACAG-----TACTCTCTGAGAAATCATTG	2751
OY	964	ThrCysleuserGIu-----TyTyraSPhleuProAsnGIuValGIuLyS	978
		
Db	2752	CCAACCTGTGGTGGAAAGAAAGACAGAGACATTTGAATCTGCACAGAGGCCCTCTCAGC	2811
OY	979	LeuLyasAn-----GIuIleSerLySAlaSerGIuGIuIleIleuIeuleuLySgIn	995
		
Db	2812	TTTAACAACAACCTCTTATTCAGTAGAAAGAAATCGAAGAA-----	2853
OY	996	GIuGIuLyGIuHIsSerAlaSerIleIleSerLySgInGIuIleIleuGIuInser	1015
		
Db	2854	-----AAGAAAGAACCTTGACACAGTACAGGCC	2883
OY	1016	GIu-----GIuIleuGIuIeuThrAspGIuValThr-----His	1027

Db	3721	---	TCCTGGCAGACAGTTCAGTACTGCTCTACACAGACAGTGAAGTCATGGGAGACT	3777
Qy	1346	LeuGlnThrSerCysIysAlaLeuVal	SerGluLeuGluLeuLeuArgAlaHisVal	1365
Db	3778	GGCCATGACCAAGTGCAGTCCAGTGCATGTC	---	CTTCTGACACCCACCATTTGCA 3825
Qy	1366	SerValGlu	-----	1368
Db	3826	CTGGTGAAGGAGACCTGTGTTTTTATCCAGCATTCATCTCGAAACATACCTCCCTCG	3885	
Qy	1369	GlyGluAsnLeuGluIleThrIly	SlyLeuAsnGlyLeuGluLys	-----Glu 1384
Db	3886	GGTGCACAAATTGATGATGCATC	---AAATGCATGCTTTAAGTAATGCAGGTGTGTT	3942
Qy	1385	IleLeuGlyIysSerGluGluSerGluValLeuLysSerMetLeuGluAsnLeuLysGlu	1404	
Db	3943	GTTCACGAAAGAAATAATGTGTCAACAGTGAACATGCTCTTCTTACAGAACTCAAACT	4002	
Qy	1405	AspAsnAsnLysLeuLysGluGlnAlaGluIlyTrpSerSerIysGluAsn	-----Gln 1422	
Db	4003	TCAGTGGGTTCAGAAATATGATCCACCTGCACACCTTCAGAGAGCCCAAAATTCACAGTGG	4062	
Qy	1423	PheSerLeuGluGluValPheSerGlySerGlnIysLeuValAspGluIleGluValLeu	1442	
Db	4063	TTCAACACCCCATTTGATCTTCTCAAGCAGTCAAAATGTGCACCT	-----GAGGTCTGG 4116	
Qy	1443	LysAlaGlnLeuLysAlaAlaGluArgLeu	1453	
Db	4117	AAACTTACTTTCATTCATTCGAAGATGAGGCTCTT	4149	
RESULT 9				
US-09-721-832-1				
; Sequence 1, Application US/09721832				
; Patent No. 6399346				
; GENERAL INFORMATION:				
; APPLICANT: Beraud, Christophe				
; APPLICANT: Freedman, Richard				
; TITLE OF INVENTION: No. 6399346el motor proteins and methods for				
; FILE REFERENCE: 1055				
; CURRENT APPLICATION NUMBER: US/09/721,832				
; CURRENT FILING DATE: 2000-11-24				
; NUMBER OF SEQ ID NOS: 4				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 1				
; LENGTH: 4176				
; TYPE: DNA				
; ORGANISM: Human				
US-09-721-832-1				
Alignment Scores:				
pred. No.: 1.65e-64 Length: 4176				
Score: 915.50 Matches: 416				
Percent Similarity: 40.59% Conservative: 246				
Best Local Similarity: 25.51% Mismatches: 536				
Query Match: 6.20% Indels: 433				
Gaps: 71				
DB: 4				
US-09-150-867-1 (1-2954) x US-09-721-832-1 (1-4176)				
Qy	6	AlaValLysValCysValArgValArgProIleuLeuGlnArgGluGlnLysAspGlnAla	25	
Db	7	TCGGTCACAGTGGCCGTAGGGGTCCGGCCCATGATGCGAGGGGAAAGGACTTGGAGGCC	66	
Qy	26	Asn-----LeuGlnTrpLysAlaGlyAsnAsnThrIleSerGluVal	39	
Db	67	AACTTCATTATTCAGATGAGAGAAAGCAAAAGCAATCAACAACCTTAAAGATACAGAA	126	
Qy	40	-----AspGlyThrIlySerPheAsnPheAspArgValPhe	51	
Db	127	GGAGGCACTGGGAGACTCGAAGAGAACGACCAAGACCTTCACTTATGACCTTTCTTTT	186	
Qy	52	AsnSerHisGluSerThrSer-----GlnIleTyrgGlnIle	64	

Oy	719	PheaspasniIlelleLeuglunsglnIuThrLeuIysArgIuIleAlaAspLeuGlnArg	738
Db	2089	-----CAAGAAAGAAAGACCTTTCCTCCCGCTCAAGAAAGAACTCCAAACA	2133
Oy	739	-----SerLeuYsglnuAsnglnIuThrAsnglnPheglnIleLeuGlnLys-----	754
Db	2134	CTCAAAAGAACTAACACACACAGAAAGGCTGAGAACTTCAGATTTTCAGAACTGGAC	2193
Oy	755	gluThrGlnLysglnuHsglnuAlaLeuIleHisIsglnIleGlySerLeuLysLysLeu	774
Db	2194	CAGCTCCAAAGAAAGAAAAGATGACAGATATGCC-----AAGCTGAACTGGAAAAAAGAGA	2250
Oy	775	ValGlnuAsnAlaGlu-----MetTyrAsnGlnAsnglnuGlnuAspLeuGlu	790
Db	2251	CTAGAGGACGACGAGAAAGACGACGATCTCTCGTCGGCCATCTGGAAAGACGCTCGCA	2310
Oy	791	ThrLys-----ThrLysLeuLeuYsglnuGlnuIleGlnLeuAlaGluLeuArg	807
Db	2311	GAGAAACGAGAGATGATCCAGCTCTCGCGCGCTGGGAGATGACAGGGGTGGAAAGAG	2370
Oy	808	LysArgAlaAspAsnLeuGlnLysLysValArgAsnPheAspLeuSerAlaSerMetGly	827
Db	2371	AAGAG-----	2376
Oy	828	AspSerGlnLysLeuCysGlnuIlePheGlnLeuLysGlnSerLeuSer-----	844
Db	2377	GACCTGGAAAGCATTCGGGGATCCCTCGCGGGTGAAGAGAGCTCTCCGAGGGGAT	2436
Oy	845	---AspAlaGlnuAlaValThrArgAspAlaGlnLysGlnCysSerPheLeuArgSerGlu	863
Db	2437	GAGATGGCGAGAGATGAGAAAAGGCTCACTCGCTGTTCTTCGATTTCAAGACAGAGCAG	2496
Oy	864	AsnLeuGlnLeuLeuYsglnuYsglnuYsglnuAspThrSerAsnTrpTyrAsnGlnLysGlnLys	883
Db	2497	CTTGTCAGCTAGTGAACCTCGAAGAAAGACCTGCTCAGACAGAAACATCTCGAAAAA	2556
Oy	884	AlaAlaSerLeuPheGlnLysGlnLeuGlnIuThrGlnLysSerAsnTyrLysLysMetGlu	903
Db	2557	GAACTCCAAAGAAAGACGAGATCTTAGAGTGTTTAAAGTGAACATGACAAAGATCT	2616
Oy	904	AlaAspLeuGlnLysGlnLeuGlnSerAlaPheAsnGlnLysAsnTyrLeuAsnGlnLeu	923
Db	2617	AGATTGTTGGAAAAACATGATGAGAGTGCACAGAT-----	2652
Oy	924	LeuAlaGlnLysValProArgAspLeuLeuSerArgValGlnLeuGlnLysLysValSer	943
Db	2653	---GTCAAGGAAAGCTCAAGATTTTCAGAAAAATAAAGCCAGTGCAGTCCAGGCTCA	2709
Oy	944	GlnPheSerLysGlnLeuGlnLysAlaLeuGlnuGlnLysAsnAlaLeuGlnuAsnGlnuAla	963
Db	2710	TATAAAGACGCCAGCTACAG-----TACCTCGTGGAAATCACTTG	2751
Oy	964	ThrCysLeuSerGlu-----TyrLysPheLeuProAsnGlnuValGlnCys	978
Db	2752	CCAACCTGTTGGAAAGAAAGACAGAGACATTTAAATTTCTTGACAGAGGCTCTCAGC	2811
Oy	979	LeuLysAsn-----GlnIleSerLysAlaSerGlnuIleuMetLeuLeuYsgln	995
Db	2812	TTAGACACACACTTTTATCAAGTGAAGAAATGAGAA-----	2853
Oy	996	GlnGlnGlnHisSerAlaSerIleIleSerYsglnuIleIleMetGlnuGlnuSer	1015
Db	2854	-----AAAGAAAGACGCTGCACAGTACCAAGGCC	2883
Oy	1016	Glu-----GlnIleLeuGlnLeuThrAspGlnuValThr-----His	1027
Db	2884	AATGCAAAACCAAGCTCAAAAGCTCCAAACCACTTTGAATTCACCTGCCACATTGCACGT	2943
Oy	1028	ThrGlnSerLysValGlnGlnThrGlnGlnuGlnuTyrLeuGlnuMetLysLysMetHisAsp	1047
Db	2944	CAGGAGAAAGAAAGTAAAGAAAGAAAGCAATTTTGGAGTCC-----	2988

QY	1048	AspLeuPheGluIysTYrTlleArgAsnLysSerGluAlaGluAspLeuLeuArgGluMet	1067
Db	2989	-----AGAGAGAAAGCAGCAGAGAGCGCGCTGAGCGGCCCTG	3027
QY	1068	GluAsnLeuIysGlyThrMetGluSerValGluValIysIleAlaAspThrLysHisGlu	1087
Db	3028	-----AGCAGCTCC	3069
QY	1088	---LeuGluIuThrIleArgAspLysGluIleLeuLeuHisGluIysTYrPhePhe	1106
Db	3070	ACCCCTGGCGACGAGATTGAGAGCAGAGCGAG	3102
QY	1107	GlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLys	1126
Db	3103	-----AAA	3105
QY	1127	LeuValGluGlyAsnSerGlnAspProIleGluIleAsnAspTYrHisAsnLeuIleAla	1146
Db	3106	CTTGCCAGCTCTCAACAGTGGC	3126
QY	1147	LeuAlaThrGluArgAsnAsnIleMetValCysLeuGluIuThrGluArgAsnSerLeuLys	1166
Db	3127	---AGCAGAGACACAGTCAGCGGCTCCAGCGTACCTGAGCGAGCGACAGAGAACCCCTGGAG	3183
QY	1167	GluGluIValIleAspLeuAsnThrGluIleuGlnSerLeuGlnAlaGlnSerIleGluLys	1186
Db	3184	ANGGACCAAGAGAGGTTAATGAAATATGAAATCCAGCAGCTGAACACAGAAATTTATGAGTGC	3243
QY	1187	SerAspLeuGlnIuThrProLysGlnAspLeuGluGluIValIleLysLeuLeuLeu	1205
Db	3244	GATGGTGTTCAAAATCATCATCATGGAGGCCCTGGAAAGGAAAGTGGCTTCCAGCTTG	3303
QY	1206	GluMetGluLeuLeuLysGlyHisLeuThrAspSerGlnLeuSer	1220
Db	3304	CCAGTCAGTCTGAAAATATCACACCTGGTCTCCCTCATGATGATCCAGATCAATCTTAC	3363
QY	1221	IleGluLysLeuGlnIleuGluAsnLeuGluIValIThrGluLysLeuGlnThrLeu	1238
Db	3364	ATTGAAAGA---GANGTCAAAAGACGCTTCAGGATTTGCATGCT	3405
QY	1239	---GlnGluIuMetLys	1243
Db	3406	GTGATTAGTGAAGCGTCACTACATCTGCACAGCATGAGATTAATGAGAACTTCAC	3465
QY	1244	AsnIleThrIleGluArgAsnGluLeuGlnThrAsnPheGluAspLys	1260
Db	3466	AATGGCACCATTTCAACGTAAACTAAATAATATGACGTGTGTGCTACTCCTGTGTCTCG	3525
QY	1261	---AlaGluHisAspSerLeuLysGlnAspLeuSer	1271
Db	3526	ATGCCAGACGCTGATGCCGTGCCGTGGCTTAATCATCCCTTGTCCAGAGAAATCTG---	3582
QY	1272	GluAsnIleGluIuThrIleGluIuThrGlnAspGluLeu---Arg	1285
Db	3583	---GTTACGCTTCTCTCTGATTGGAAAAACGAAATCCCTGATTTAGTTTGCCAAAT	3636
QY	1286	AlaIaIaGlnGluIuLeuArgGluGluIuLysGlnIleuValAspSerPheArgGlnIuLeu	1305
Db	3637	GGAGTTCAAGGTGTCATCCAAATTCACACATCACTGTGTGACATGAT---TACTTT	3690
QY	1306	LeuAspCysSerValGlyIleSerSerProAsnHisAspAlaValAlaAsnGlnIuLys	1325
Db	3691	CTTCATGGAATATGAGAATCAATGTCCT---	3720
QY	1326	ValSerLeuGlyGluValAsnSerLeuGlnSerGluMetLeuArgGlyIuArgAspGlu	1345
Db	3721	---TCCCTGGCAGAGATTCACTGTCTTACACAAACAGTGAAGTACGGGGTCTCT	3777
QY	1346	LeuGlnThrSerCysLysAlaLeuValSerGluLeuGluIuLeuLeuArgAlaHisValLys	1365
Db	3778	GGCATTGACCAAGTGCACATGCTGATGC---CTTCGAACACCCACATTGCA	3825
QY	1366	SerValGlu---	1368

```

|||||
Db 3826 CTGGTGAAGAAAGCTGTTTATCCAGCATTCGATCGAAGATACCTCCTCG 3885
Qy 1369 GtGuaenleuGluileThrlLysLysLeuansnglyLeuGluLys-----Glu 1384
Db 3886 GGTGCACAAATTTGAGTGATC--AAATGCCATGCTTTAAGTAATTCAGGTGTGTT 3942
Qy 1385 IleuGluLysSerGluLysSerGluValLeuLysSerMetLeuGluValLeuLysGlu 1404
Db 3943 GTTCCAGAAAAAGAAAGTGTCAACAGTACAGTACTGCTTCTTACAGAACTCAAACT 4002
Qy 1405 AspAsnAsnLysLeuLysGluLysGluLysGluLysSerSerLysGluAsn-----Gln 1422
Db 4003 TCAGTGGGTTCCAGAAATAGTCCACCTGACACCTTCAGAAAGCCCAATGTCAGTTG 4062
Qy 1423 PheSerLeuGluLysValPheSerGlySerGluLysLeuValAspGluIleGluValLeu 1442
Db 4063 TTTCACCAACCCCATGTTATCTTCAAGGCAGTCAGATGTCGACCT-----GAGGTCTGG 4116
Qy 1443 LysAlaGluLeuLysAlaAlaGluLysGluLysGlu 1453
Db 4117 AAATCTACTTCATCTCAAGATGAGCTCTT 4149

RESULT 10
US-09-721-689-1
; Sequence 1, Application US/09721689
; Patent No. 6440685
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6440685el motor proteins and methods for
; FILE REFERENCE: 1055
; CURRENT FILING DATE: 2000-11-24
; CURRENT FILING DATE: 2000-11-24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4176
; TYPE: DNA
; ORGANISM: Human
US-09-721-689-1

Alignment Scores:
Pred. No.: 1.65e-64 Length: 4176
Score: 915.50 Matches: 416
Percent Similarity: 40.598 Conservative: 246
Best Local Similarity: 25.518 Mismatches: 536
Query Match: 6.208 Indels: 433
Gaps: 71

US-09-150-867-1 (1-2954) x US-09-721-689-1 (1-4176)
Qy 6 AlaValLysValCysValArgValArgProLeuIleGlnArgGluGluLysAspGlnAla 25
Db 7 TCAGTGTGAGTGGCGGCGGAGGTCGCGCCATGATCCAGAGAAAGAGACTTGGAGGCC 66
Qy 26 Asn-----LeuGluIlePlyLysAlaGlyLysAsnThrIleSerGluVal----- 39
Db 67 AAGTTCATTATTCAGATGAGAGAAAAAGCAAAAGCAATCAAACTTAAAGATACCCAGAA 126
Qy 40 -----AspGlyThrLysSerPheAsnPheAspArgValPhe 51
Db 127 GAGGACACTGGGAGCTCAGAGAGAAAGCAAGACCTTACACTTATGACTTTCTTTT 186
Qy 52 AsnSerHisGluSerThrSer-----GlnIleTyrgLysGluIle 64
Db 187 TATTCCTGCTATCAAAAAAGCCAGATTACGTTTCACAGAAAGATGCTTTTCAAAACCTTC 246
Qy 65 AlaValProIleIleArgSerAlaLeuGluGlyTyraSnglyThrIlePheAlaTyrgly 84
Db 247 GGCACAGATGTCGTAAGTGTGCAFTTGAAGGTATATATGCTTGTGTTCATATGGG 306

```

```

Qy 85 GlnThrSerSerGlyLysThrTyrrhMetMetGlyThrProAsnSerLeuGlyIleIle 104
Db 307 CAAACTGATCTCGAAGAGCATACACTATGATGGGAATTCGAGATTTGCTTAATA 366
Qy 105 ProGlnAlaIleGluGluValPheLysIleIleGluIle-----ProAsnArg 121
Db 367 CCTCGAGATCTGTGAAGAGCTCTTCAGTCGATTAATGAACACAGATGGATGAGACT 426
Qy 122 GluPheLeuLeuArgValSerTyrrhMetGluIleTyraSnglyThrValLysAspLeu 141
Db 427 TCTTTTCGAACTGAAGTACGCTACTTAGAAATTTATACGAAGCTGTGAGATGACTACT 486
Qy 142 CysAspAspArgArgLysLysPro-----LeuGluIleArgGluAspPheAsn 157
Db 487 -----CGCGGAGAGCATCTAAACCTTCATTTGAGAGTCCGTAGAGATCCCAA 537
Qy 158 ArgAsnValTyrrhAlaAlaAspLeuThrGluLeuValMetValProGluIleValIle 177
Db 538 GAAGCCCTTATGTGAGGATTTATCCAAACATTTAGTACAGAAATATGTCAGCTAGAA 597
Qy 178 GlnTPriLeuLysGlyGluLysAsnArgHisTyrglyGluThrLysMetLysAsnHis 197
Db 598 GAACCTTAGGATCGGGCATATATCAACCGACCGACCGACGCTGGGTGAACAGACGTC 657
Qy 198 SerSerArgSerHisThrIlePhe-----ArgMetIleValIleGluSerArgAspArgAsn 215
Db 658 ACTAGCAGGTCTCATGCCATCTTCAACATCAAGTTCACACGAGCTTAATTTATTTCTGAA 717
Qy 216 AspProThrAsnSerGluAsnCysAspGlyAlaValMetValSerHisLeuAsnVal 235
Db 718 ATGCCA-----TGTGAA-----ACCGTCAGTAAGATCCACTGGTT 753
Qy 236 AspLeuAlaGlySerGluArgAlaSerGlnThrGlyAlaGluLysValArgLeuLysGlu 255
Db 754 GATCTTCGCGGAGAGTGAAGGTGATGATGCCACCGAGACCGGGGTATGCTTAAGGA 813
Qy 256 GlyCysAsnIleAsnArgSerLeuPheIleLeuGlyGluValIleLysLysLeuSerAsp 275
Db 814 GGGGAAATATATACAAAGATCCCTTGTGACTCGGGGAAGCTGATTTTCGCTTACTGAT 873
Qy 276 GlyGlnAlaGlyGly-----PheIleAsnTyrrg 285
Db 874 TTATCTCAGATGCTCAAAATCTCTTGCMAAGAAAGCAAGTTCGTGCTTACAGG 933
Qy 286 AspSerLysLeuThrArgIleLeuGlnAsnSerLeuGlyLysAlaLysThrValIle 305
Db 934 GATTTGTGTTGACTTGGTGTGTTAAAGTACCTTGGGGAACCTTAACCTATTCATG 993
Qy 306 IleCysThrIleThrPro-----ValSerPheAspGluThrLeuSerThrLeuGlnPhe 323
Db 994 ATTGCCACCATTTCACTCGCTGATGTCAATTATGAGAAACCTTAAGTACTTGGCAT 1053
Qy 324 AlaSerThrAlaLysHisValArgAsnThrProHisValAsnGluValLeuAspAspGlu 343
Db 1054 GCAAAATAGAGCCAAAAACATCATCAACAAGCTTACATTAATGAGATGCCAACCTCAAA 1113
Qy 344 AlaLeuLysLysArgTyrrgLysGluIleLeuAspLeuLys----- 357
Db 1114 ---CTTATCCGAGACTCGAGCTGAATATAGCCAGACTGAAACGCTGCTGCTCAAGG 1170
Qy 358 LysGlnLeuGluAsnLeuGluSerSerSerGluThrLysAlaGlnAlaMetAlaLysGlu 377
Db 1171 AATAGATATGGCCCTTACACTCCGCCCA-----GCTTAAGTATGAG 1215
Qy 378 GluHisThrGlnLeu-----LeuAlaGluIleLysGlnLeuHisLysGluArgGluAspArg 396
Db 1216 GAAAAAATTCAGCAAGATGAAGCAAGAGCTTCAGAAATGACCAAGATGACAAATAG 1275
Qy 397 IleThrPheLysLeuThrAsnIleValValAlaSerSerGlnGluSerGlnAlaAspGlnArg 416
Db 1276 ---TGAATGAAC-----CAAAAT 1293
Qy 417 ValLysArgLysArgArgValThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAla 436

```

```

Db 1294 ATTTGAAAGAACAACTCTA-----GCCCTCAGGAAA 1326
Qy 437 SerGlyValSerAspPheAspMetLeuSerArgLeuProGlyAsnPheSerIysVala 456
Db 1327 GAAGGGATTGGA-----GTTGTT 1344
Qy 457 LysPheSerAspMetProSerPheProGluIleAspAspSerValGlyThr----- 473
Db 1345 TTGATATTCGAACTGCTTATTCATTTGATTCGATGATGATCCTTTGATGACTGGAATC 1404
Qy 474 -----GluPheSerAspPheAspAspAlaLeuSerMet 484
Db 1405 ATCTTATATCATTTAAAGAGAGTCACACATACGTTGCTGAGACACATGCTTCCACGGAG 1464
Qy 485 MetAsp-----SerAsnGlyIleAspAlaGlu-----TrpAsnLeu 496
Db 1465 CAAAGATATTCCTTCCTGACCTTGACCTTGAGAGAGTGCATTCATCTTGAATAATATC 1524
Qy 497 AlaSerIysValThrHisArgGlyIleThrSerLeuHisGlnSerMetIleAspPheGly 516
Db 1525 GGGGGGACAGTACT--CTGATACCCCTGAGTGGTCCAGTGTCTGTGAATGCTGT 1581
Qy 517 GlnIleSerAspSerValGlnPheHisAspSerSer-----LysGluAsn 531
Db 1582 CAGATCTGGAGGCGCACACATCTAAATCAAGGTGCTGTGATTCCTTGGAGAACCAAT 1641
Qy 532 GlnLeuGlnIyr--LeuProLysAspSerGlyAspMetAlaGlySerArgLys----- 548
Db 1642 ATGTTTCGCTTAAACCTCAACAAAGAGACCGCAAGCTCAGGGAAGAGAGAGTGC 1701
Qy 549 -----AlaSerPheGluIysGluIleThrSerLeuGlnGlnLeuGlnSerIysGlu 566
Db 1702 CTTCGTGCTCTCTCCTCAGCTGTGCTCATGACCGACCTC-----TCGAATCC 1746
Qy 567 GlnGluIysIysGluLeuValGln-----SerPheGluLeuIysIle 580
Db 1747 CGTGAGAAACCTGTCTGCATGCTGTATTAACCCGAGCTTGAATTTGAGAGCAACAG 1806
Qy 581 AlaGluLeuGlnGlnIleuSerValIysAlaIysAsnLeuGlnMetValThrAsnSer 600
Db 1807 CGTAAAGAACTTGAATAATTGAAAATTAAGAACTCAATAGAAAGAAATG--GAGGAA 1863
Qy 601 ArgGlnHisSerIleAsnAlaGluValGln--ThrAspValGlnGluValValArg 619
Db 1864 AAGCAGAAATCAGACAGGCGTGAAGCTGAGCGATGACGACGAGAGTGGAGAACCCAGCG 1923
Qy 620 LysGlnMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAsp 639
Db 1924 AAGGAGACAGAAATCGTG-----CACTCCAG-- 1950
Qy 640 SerSerValAspGlyLysArgLeuSerSerSerHisAspGlyCysIleGlnHisArgLys 659
Db 1951 -----ATTCGACAGCAGAGAGAGGACCTCAAAAGCGCGCAGC 1986
Qy 660 Met--LeuGlnGlnIysIleValAspLeuGlnGluPheIleGluAsnLeuAsnLysLys 678
Db 1987 TTCACATCGACAGAACCTTAAGGATTTA-----CTTGGCGAGAG 2028
Qy 679 SerGluAsnAspLysGlnLysSerSerGlnLysPheMetGluSerIleGlnLeuLys 698
Db 2029 GAAAAAATTTGAAGAGGAGGCTGAGGAGCAG----- 2061
Qy 699 GluAlaIleMetAlaGlnLysAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAspAsn 718
Db 2062 -----CAGGAAATCGAGCTGCAGAAAGAGAGA 2088
Qy 719 PheAspAsnIleIleLeuGlnAsnGluThrLeuLysArgGluIleAlaAspLeuGlnArg 738
Db 2089 -----CAAGAGAGAGAGACTTCTCCGCTCCAGAGAACTCCAAACA 2133
Qy 739 -----SerLeuLysGluAsnGlnGluThrAsnGlnPheGluIleLeuGlnLys----- 754

```

```

Db 2134 CTCAAAGAACTCAACACAGAGAGGCTGAGAACTTCAATATTCAGAACTGAGAC 2193
Qy 755 GluThrGlnLysGlnHisGluIleGlnLeuIleHisGluIleGlySerLeuLysLysLeu 774
Db 2194 CAGCTCCAAAGAGAAAAAGATGACAGATATGCC--AAGCTTGAACCTGAGAAAAAGAGA 2250
Qy 775 ValGluAsnAlaGlu-----MetTyrAsnGlnAsnLeuGlnGluAspLeuGlu 790
Db 2251 CTAGAGAGCAGACAGAGAGAGAGAGCATGCTCGTGGCCCATCTGGAAGAGAGCTCCGA 2310
Qy 791 ThrLys-----ThrLysLeuLeuLysGlnGlnGlnIleGlnLeuAlaGluLeuArg 807
Db 2311 GAGAAAGCAGAGATGATCCAGCTCCTCGCGCTGGGAGAGTACATGGGTGAGAGAGAG 2370
Qy 808 LysArgAlaAspAsnLeuGlnLysValArgAsnPheAspLeuSerValSerMetGly 827
Db 2371 AAGAG----- 2376
Qy 828 AspSerGlnLysLeuCysGlnGluIlePheGlnLeuLysGlnSerLeuSer----- 844
Db 2377 GACCTCGAAGCAGATTCGGGAATCCCTCGCGGTGAGAGAGCTTCGCGAGAGGAT 2436
Qy 845 -----AspAlaGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerGlu 863
Db 2437 GAAGATGGCGAGAGTACAAAAGGCTCAACTGCTGTTCTTGCAATTCAGAGAACCGAG 2496
Qy 864 AsnLeuGluLeuLysGlnLysMetGluAspThrSerAsnTrpTyrAsnGlnLysGluLys 883
Db 2497 CTTCGCAAGCTATGAACTTGAGAGAGACCTGTTCAACAGAAAGACATCTCGAAGAAA 2556
Qy 884 AlaAlaSerLeuPheGlnLysGlnLeuGlnIleThrGlnLysSerAsnTrpLysLysMetGlu 903
Db 2557 GAAGTCCAGAGAGAGAGAGATCTAGAGTGTAAATGTGAACTATACAAAGAAATCT 2616
Qy 904 AlaAspLeuGlnIysGlnLeuGlnSerAlaPheAsnIleAsnTrpLeuAsnGlnLeu 923
Db 2617 AGATTGTTGAAAACATGATGAGCTGCACAGAT----- 2652
Qy 924 LeuAlaGlyValProArgAspLeuLeuSerArgValGlnLeuGlnLysValSer 943
Db 2653 ---GTCACGGAAGTGCCTCAAGATTTTCAGAAAATTAAGCCAACTGAGTACAGGCTGCA 2709
Qy 944 GluPheSerIysGlnLeuGlnLysAlaLeuGlnLysAlaLeuGlnAsnLeuGlnVal 963
Db 2710 TATTAAGAACGCCAGCTACAG-----TACTCTCGCAGAAATCTCTG 2751
Qy 964 ThrCysLeuSerGlu-----TyrLysPheLeuProAsnGluValGluCys 978
Db 2752 CCAACTCTGTTGAGAAAAGCAGAGAGCATTTGAATCTTGACAGAGGCTCTCAGC 2811
Qy 979 LeuLysAsn-----GlnIleSerLysAlaSerGlnGluIleMetLeuLeuLysGln 995
Db 2812 TTAGACAAACACTTATCAAGTAAAGAAAAGAAATGGAATA----- 2853
Qy 996 GluGlyGlnHisSerAlaSerIleIleSerLysGlnGluIleIleMetGlnGlnIns 1015
Db 2854 -----AAAGAAAGAACGTTGCAACATGACATGACAGGCC 2883
Qy 1016 Glu-----GlnIleLeuGlnLeuThrAspGluValThr-----His 1027
Db 2884 AATGCAAAACAGCTCCAAAGCTCAAGCCACTTGAAATTCCTGCACACATTCGACGT 2943
Qy 1028 ThrGlnSerLysValGlnGlnIleThrGlnGlnIleTyrLeuGlnMetLysLysMetHisAsp 1047
Db 2944 CAGAGAGAAAAGTGAAGAAAAGGAAAGAGATTTTGGAGTCC----- 2988
Qy 1048 AspLeuPheGluLysTyrIleArgAsnLysSerGlnAlaGlnAspLeuLeuArgGluMet 1067
Db 2989 -----NAGAGAACACAGAGAGAGAGCGCTGAGAGCGGCGCTG 3027
Qy 1068 GluAsnLeuLysGlyThrMetGluSerValGluValLysIleAlaAspThrLysHisGlu 1087
Db 3028 GCCAGGCTGAGAGAGACATTCGCGCTCAG-----AGGCACCTCC 3069

```


Dd	430	AAAGAAATTGATGAAAAAGAGTAGCTATTGATTTACTCGAAGAGTGCTTACTTACGAAATT	489
Qy	133	TyrAnGIuThrValLysAspLeuEncyCysAspArgLysProLeuGIuIle	152
Dd	490	TACATGCAAAATTTTGGATCTTCATGTCATGTCATGTCGAGAAAGTCATTAGAGATT	549
Qy	153	ArgGIuAspPheAsnArgAsnValLysValAlaAspLeuThrGIuGIuLeuAlaMetVal	172
Dd	550	TACATGCAAAATTTTGGATCTTCATGTCATGTCATGTCGAGAAAGTCATTAGAGATT	609
Qy	173	ProGIuHISValIleGIuInTrIleLysLysGIuLysAsnArgHISLysLysGIuThr	192
Dd	610	GCGTTGGATCTTCTTCCTGTTTGGAAACAGCAACACTGAGACTGAGGCTCAGC	669
Qy	193	LysMetAsnAspHisSerSerArgSerHISThrIlePheArgMetIleValGIuSerArg	212
Dd	670	GCTATGACATCCCGACGCGCCGATCTATGCAATTTAACAACTCTTTAGGCAAGGA	729
Qy	213	AspArgAsnAspProThrAsnSerGIuAsnLysAspGIuAlaValaMetAlaSerHISLeu	232
Dd	730	AGAGAAAGTGACAAAGATAGCAGC-----TTGGCTCCAAGCTG	768
Qy	233	AsnLeuValAspLeuAlaGIuSerGIuArgAlaSerGIuThrGIuAlaGIuLysValArg	252
Dd	769	CATCTTGTAAGCTCGCTGGATCAGAAAGACAGAAAGAAACCAAGCTGAGGGAGTCGT	828
Qy	253	LeuLysGIuGIuCysAsnIleAsnArgSerLeuPheIleLeuGIuGIuValIleLysLys	272
Dd	829	CTAAAGAGGGTATTATTATTACCGAGGCTCTCATGCTTGGAAATGTAATCACTGCT	888
Qy	273	LeuSerAspGIuAlaGIuLysIleLysIleAsnLysArgAspSerLysLeuThrArgIle	292
Dd	889	CTTGAGATGACAAAGAGGGTGCTTGCGCCCTACAGAGATTCAGAGTTCAAGTTGACTGACG	948
Qy	293	LeuGIuAsnSerLeuGIuLysAlaAsnLysThrValIleLeuSerHISThrIleThrProVal	312
Dd	949	CTTCAAGATTCCTCGAGAGGTATTATGCAATCTGTATTATGATGCTGTGTGAACTCTGCT	1008
Qy	313	-----SerPheAspGIuThrLeuSerThrLeuGIuAlaPheAlaSerHISAlaLysHISVal	330
Dd	1009	GACTCCAAATCTGAGAGAAACATTAAATACCTTGCTGATCTGACAGAGCAAGAAATAC	1066
Qy	331	ArgAsnThrProHISValAsnGIuValaLeuAspArgAspGIuAlaLeuLeuLysArgLysArg	350
Dd	1069	AGAGAAATACCTATGTTAAAT-----ATTGATCCCGACACA-----	1104
Qy	351	LysGIuIleLeuAspLeuLysGIuLeuGIuAsnLysLeuSerSerGIuThrLys	370
Dd	1105	GCTGAACCTTAATCACTTAACCAACAGAGTCAACACAGCTA-----	1143
Qy	371	AlaGIuAlaMetAlaLysGIuGIuHISThrGIuLeuAlaLeuGIuLysGIuLeuHIS	390
Dd	1144	-----CAAGCTGTGTTG-----	1155
Qy	391	LysGIuArgGIuAspArgLIEtrPheLysLeuThrAsnIleValAlaLaseSerGIuGIu	410
Dd	1155	-----	1155
Qy	411	SerGIuGIuAspGIuArgValLysArgLysArgValaThrTrpAlaProGIuLysIle	430
Dd	1155	-----	1155
Qy	431	GIuAsnSerLeuHISAlaSerGIuValaSerAspPheAspMetLeuSerArgLeuProLys	450
Dd	1156	-----CTACAGGCCCATGGA-----GTTACCTCGCTCGGA	1185
Qy	451	AsnPheSerLysLysAlaLysPheSerAspMetProSerPheProGIuIleAspAspSer	470
Dd	1186	TCTTAATACTGTGGA-----CCATCA-----	1206
Qy	471	ValCysThrGIuPheSerAspPheAspAlaLeuSerMetAspSerAsnGIuLysIle	490

Dd	1207	-----GAGAACTCAACATCCCTGATGGAGAAAT-----	12386
Qy	491	AspAlaGluTyrAsnIleuAlaSerLysValThrHisArgLysThrSer-----Leu	508
Dd	1237	-----CACTCCCTGGTAAAGAGAAATATAAAATTAAGTGTGTGCTG	1278
Qy	509	HisGlnSerMetIleAspPheGlnIleLeuSerAspSerValGlnIleHisAspSerSer	528
Dd	1279	AGCGAGGCGAGCTGTCAACAGACCACAGATGTGTGGAGGATCTTTGGACAGAGCAAGCCG	13388
Qy	529	LysGluAsnGlnIleuGlnTyrLeuProLysAspSerGlnLysPheIleLysArgLys	548
Dd	1339	AATGAA-----AAATGAACCCCAAGCTTAAGAGCTCGAGGACG	13777
Qy	549	--AlaSerPheGlnLysGlnIleThrSerLeuGlnGlnIleuGlnIleuGlnSerLysGlu--	566
Dd	1378	CAGCGCGCTGCAAACTGATCTTCAAAAGCTAGTGGACACTTTGGAAAGCCAGAAATTG	14378
Qy	567	GlnGluLysLysGlnIleuValGlnSerPheGlnIleuLysIleAlaGlnIleuGlnIleu	586
Dd	1438	AAAGAAATGTAGAAATATTGTGAACCTGACAGATTGTATCCAGATTACCGATTACGGATAA	14978
Qy	587	LeuSerValLysAlaLysAsnIleuGlnMetValThrAsnSerArgLysIleSerIleAsn	606
Dd	1498	-----ACTGTTGCTGCATGGCTGACGACCATTTGATCTCGGTGAG	15339
Qy	607	AlaGluValGlnThrAspValGlnLysGlnValValArgLysGlnMetSerValIleuGly	626
Dd	1540	CAAGAGCCCAAGTAGAAACCATGCCAGAG-----	15669
Qy	627	AspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSerSerValAspLysArg	646
Dd	1570	-----ACGAGAGGCTTCTTGACGCGTTTACACATCAGCATGCTCTCGT	16144
Qy	647	LeuSerSerThrAspGluCysIleGlnHisArgLysMetLeuGlnIleLysIleVal	666
Dd	1615	CAAGGCGAGATGTGTAAAGACCTGGTGAATTAAGACCCCTTGACCTGCACTGAA-----	16688
Qy	667	AspLeuGlnIleuPheIleGlnAsnIleuAsnLysLys-----SerGlnAsnArgGlnIle	684
Dd	1669	-----GAGGCGCTGGCAGAGAAATACATCAGAAATGACAGCCCACTG	17110
Qy	685	LysSerSerGlnIleuAspPheMetGlnSerIleGlnLeuCysGlnAlaIleMetAlaGlu	704
Dd	1711	CAGCCTATTCACTACCAATACCAAGATATACAT-----	1743
Qy	705	LysAlaAsnAlaLeuGlnIleuAlaIleMetArgAsnAsnPheAspAsnIleIleu	724
Dd	1744	AAAGAGCCAGAAATAGAAAGTCATCATCTGCSAAAGAGAAAG-----GAAAGATGGTCTT	1800
Qy	725	GluAsnGlnThrLeuLysArgGlnIleAlaAspLeuGlnLysSerLeuLysGlnIleu	744
Dd	1801	GAACTTCAAGACAGAAAGAGATGCCCAAGCCCAAGTTGATGAGAGCGCCGCCGCAAA	1860
Qy	745	GlnThrAsnGlnPheGlnIleLeuGlnLysGlnThrGlnIleGlnHisGlnAlaGlnIleu	764
Dd	1861	CGTCCCAAGAGACCTGAGAGGCTCAAAATGCTGATCGAAGAAATA-----	1905
Qy	765	IleHisGlnIleGlySerLeuLysLysLeuValGluAsnAlaGluMetLysArgGlnAsn	784
Dd	1906	CTGAATGAGCAGTCCAAATCTTGAAATTAAGAAATGCAAGAGGCTAGCTGTCCAA	1965
Qy	785	LeuGlnIleuAspLeuGlnIleuThrLysThrLysIleLeuLysGlnIleGlnIleuAla	804
Dd	1966	CTGAACCAAGAGATA-----CGAGTGAAGAAAAACCAAGCGGGGTCAAGTTA--	2010
Qy	805	GluLeuArgLysArgAlaAspPheLeuGlnLysValArgAsnPheAspLeuSerVal	824
Dd	2011	---ATGCGTCAATGAAGAAAGATGCTGAG---AAGTTTAACACTGG-----	2052
Qy	825	SerMetLysAspSerGlnLysLeuGlnIlePheGlnIleuLysGlnSerLeuSer	844
Dd	2053	-----AAGCAAAAAGAGCAAAAGAGTAATATACAGTTAAAAAGAA-----	2091

STATE: Utah
 COUNTRY: USA
 ZIP: 84070
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
 COMPUTER: AS7 Ascentia 900N
 OPERATING SYSTEM: DOS 6.22
 SOFTWARE: Word Perfect 6.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/713, 815A
 FILING DATE: 13-SEP-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Alan J. Howarth
 REGISTRATION NUMBER: 36,553
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (801)566-6633
 TELEFAX: (801)566-0750
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3572 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-713-815A-2

Alignment Scores:
 Pred. No.: 1.03e-60 Length: 3572
 Score: 867.50 Matches: 310
 Percent Similarity: 44.00% Conservative: 192
 Best Local Similarity: 27.17% Mismatches: 365
 Query Match: 5.87% Indels: 275
 Gaps: 40

US-09-150-867-1 (1-2954) x US-08-713-815A-2 (1-3572)

QY 5 AspaIaValIysValIcysValaIargValaIargProleuIleaglInglInglIaspGln 24
 DB 351 GACAGCATCAAAAGTGTCTGCCGATTCGACCGCTGAACGACGACGAAGAACGCCGCC 410
 QY 25 AlaasnleuGlntrpIysAlaIglYAsnAsn-----ThrIleSerGlnValAspGly 41
 DB 411 TCCAGTTCGTGTCACAGTCCCAACACATGTGAGAGAGAACTGCATATTCATATGCGGC 470
 QY 42 ThrIysSerPheasnPheaspArgValIpheasnSerHisGlnSerThrSerGlnIleTyr 61
 DB 471 ---AAGGTATTTCTTTCGACAAAGCTCTTCAACCGAATGCATCCAGGAAAGTCTAC 527
 QY 62 GlnGluIleAlaValIProIleIleIleArgSerAlaLeuGlnGlyTyrAsnGlyThrIlePhe 81
 DB 528 AATAGAGCGCGCAAGTCATGTATTACGAGATGCTCGCGCGGTACAAATGCAACGATATTC 587
 QY 82 AlaIyrGlnIntrpSerSerGlyLysThrTyrThrMet-----MetGlyThrPro 98
 DB 588 GCATATGTGTAGAGCTCTCCGAAAAACGCATACGATGAGGCGCGATGCGGCGACTCC 647
 QY 99 AsnSerleuGlyIleIleProGlnAlaIleGlnGlnValIlePheLysIleIleGlnGlnIle 118
 DB 648 GTAAACAGGTTATCATACACGCTATGCTCAACGACATTTTCATCATCATATGCGCATG 707
 QY 119 Pro---AsnArgIuPheLeuLeuArgValSerTyrMetGluIleTyrAsnGluThrVal 137
 DB 708 GAGGTGAACCTCGAGTTCATCATACAGTCCCTACTACGAGATCTACATGACAAAGATT 767
 QY 138 LysAspIleuLeuLysAspAspArgArgLysLysProleuGlnIleAlaGluAspPheasn 157
 DB 768 CGAATATCTGTG-----GACGTCTCCAAAGGTGAACCTTATGTGTGCAACGAGAAAGAAC 821
 QY 158 ArgAsnValIyrValAlaAspLeuThrGlnGluLeuValMetValIProGlnHisValIle 177
 DB 822 CGGGTGGCGGTACGTCACAGGCGGCTACGGAACGATTCGTCTGTCGCGCGGAGAGATGTTTC 881

QY 178 GlnTrpIleLysGlyGlnGlyLysAsnArgHisTyrGlyGluThrLysMetAsnAspHis 197
 DB 882 GAGGTGATTCAGAGAGGCAATCCAAATCGTCACATGCGCTGTACAAACATGAAACGAGCAT 941
 QY 198 SerSerArgSerHisThrIlePheArgMetIleValGlnSerArgAspArgAsnAspPro 217
 DB 942 TCTTCGGATCCCACTCAGATATTCCTATCAATGATGAG----- 980
 QY 218 ThrAsnSerGlnLysncysAspGlyAlaValMetValSer---HisLeuAsnLeuValAsp 236
 DB 981 -----CAGAGAACTCGGAGAACCCAGAACAACTATCCGGCAAACTCTACTGTGTGAT 1034
 QY 237 LeuAlaGlySerGluArgAlaSerGlnThrGlyAlaGlnGlyValArgLeuLysGlnGly 256
 DB 1035 TTGGCGGTTCCGAGAAAGTTTCCAAAGACTGAGCGGAGGAACCGTTCTGTATGAAGCC 1094
 QY 257 CysAsnIleAsnArgSerLeuPheIleLeuGlnValIleLysLysLeuSerAspGly 276
 DB 1095 AAGAACATCAACAAAGTGTGCGGCCCTTGGCAACGTAAATTCGTGCGCGGAGCGGA 1154
 QY 277 GlnAlaGlyGlyPheIleAsnTyrArgAspSerLysLeuThrArgIleLeuGlnAsnSer 296
 DB 1155 AACAA---ACGCACATCCCTTACCGTGATTCACAGTAAACGCGCATCTCGCAGAGTCG 1211
 QY 297 LeuGlyGlyAsnAlaLysThrValIleIleCysThrIleThrProValSerPheAsp--- 315
 DB 1212 CTGGGAGGCAACGACGACACACCATCATGCTGCTGCTGCTCCAGCGACTTTCACAGAA 1271
 QY 316 ---GluThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrPro 334
 DB 1272 TCTGAAACGAAAGTCACAGCTGACCTTGGTCGTAGACCCCAAGACAGTGAAGCTGTGTC 1331
 QY 335 HisValAsnGluValLeuAspAspGluAlaLeuLeuLysArgTyrArgLysGluIleLeu 354
 DB 1332 TGGCTTACAGAGAGAGCTTACTGCGGAGATGGAAGCGAGCATGATAAAGAGG----- 1385
 QY 355 AspLeuLysGlnLeuGlnAsnLeuGlnSerSerSerGlnThrLysAlaGlnAlaMet 374
 DB 1385 ----- 1385
 QY 375 AlaLysGlnGlnHisThrGlnLeuLeuAlaGluIleLysGlnLeuHisLysGlu----- 392
 DB 1386 ---AAGAAAGAAAGAGCCCACTTAAGAGGTAAAGGTGAGAGAGCTGAGATCGACTTCGC 1442
 QY 393 -----ArgGluAspArgIleThrPheIleThrAsnIle 403
 DB 1443 CGCTGAGACCGGTGAACCTGTTAAGCGGAGAGCAAAATC---AACATGAGAGATCTC 1499
 QY 404 ValValAlaSerSerGlnGlnSerGlnGlnAsp-GlnArgValLysArgLysArgValGly 423
 DB 1500 ATGAGGCAAGACAGCCCAACTGGAAGTGAGGACACACAGCGCGCGCGGAGGCC 1559
 QY 423 LThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAlaSerGlyValSerAspPheas 443
 DB 1560 GCTTTGGCGCCCAAGCA----- 1578
 QY 443 PmetLeuSerArgLeuProGlnAsnPheSerLysLysAlaLysPheSerAspMetProse 463
 DB 1579 -----CGGCTTTC-CCCAATATGTCGCGATCGGTGCGGTGACGAGCGAGCGCAG 1627
 QY 463 rPheProGlnIleAspAspSerValCysThrGlnPheSerAspPheAspAspAlaLeuSe 483
 DB 1628 GCTGGCTACAGAGTGCAGGCTCTTACACAGCGCTGAGCAAGAGGATAG----- 1679
 QY 483 rMetMetAspSerAsnGlyLysAspAlaGluTrpAsnLeuAlaSerLysValIThrHisAT 503
 DB 1679 ----- 1679
 QY 503 gGluLysThrSerLeuHisGlnSerMetIleAspPheGlyGlnIleSerAspSerValG1 533
 DB 1680 -----GAGATCAATCAGCAGAGACCA 1699
 QY 523 nPheHisAspSerSerLysGlnAsnGlnLeuGlnTyrLeuProLysAspSerGlyAspMe 543

```

Db 1700 GTACCCAGAGCTCAAGAGAGAGTGTAGAG-----CAGAGAGAACTCAT 1747
QY 543 TAlaGlucysArgLys-----AlaSerPheGluLysGluIleThrSerLeuGluGln 561
Db 1748 CGCTAACCTCGGCTGAGTGTGAGACTTTCAGTGGAGGCGCCGAATCCACAGCA 1807
QY 561 nLeuGlnSerLysGluGluLysLysGluIleValGlnSerPheGluIleAla 581
Db 1808 GAGACAGTCCGCAAGAAAGAGGTAAAGAGAGTGTCCAGAGTCCGAA----- 1856
QY 561 aGluLeuGluGluGluLeuSerValLysAlaLysAsnLeuGluMetValThrAsnSerAr 601
Db 1857 -GAGGTGACTGTAAACTACGACCAAGAAATCCAGAGAGATCAT----- 1898
QY 601 gGluHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValArgLysG 621
Db 1899 -----AACAGAAACAGAGATATGATGCTCCACAGAGAGAGCTGCACACAGA 1945
QY 621 uMetSerValLeuGluLysAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAsp 639
Db 1946 GCAGTCTGTG-----TTCACGCCGCTCCACAGAGAGTAAACAGACTCAA 1990
QY 640 -----SerSerValaAspGlyLysArgLysSerSerSerHisAspGluLysIleGluH 657
Db 1991 TGACATGCTCTCACACAGAAAGAGCGATCAG----- 2024
QY 657 sArgLysMetLeuGluGluLysIleValAspLeuGluGluPheIleGluAsnLeuAsn 677
Db 2025 -----GAAATGCTAACCAACCTACTCCGCGAGCTCGGCCAAGTGGGCCAGCCATTGCC 2080
QY 677 sLysSerGluAsnAspLysGluLysSerSerGluGlnAspPheMetGluSerIleGlu 697
Db 2081 CGCGAGTCCAGCATCGACCTTAAGATGAGT----- 2111
QY 697 uCysGluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGlu----- 711
Db 2112 -----GCTGTGCTGCGACAGGATGCCAGAGGTGAGATTTCCACCATGGCCG 2164
QY 712 -----LeuAlaLeuMetArgAspAsnPheAspAsnIleLeuGluAsnGluThrLe 729
Db 2165 TTGTGTTATTCAGCAGATGAAGAGCGAGGCCAAGAACATTGCCCGCATGCTCCACAT 2224
QY 729 uLysArgGluIleAlaAspLeuGluArgSerLeuLysGlu-----AsnGlnGluThrAsn 748
Db 2225 GGAACACAGCAGAGCTGACTCCACACAGAGATCTCCGAATATGAGAAAGATCTGGGGA 2284
QY 748 uPheGluIleLeuGluLysGluThrGlnLysGluHisGluAlaGluIleHisGluI 768
Db 2285 GTACCGGCTACTC-----ATTTCGACGACGAGCGACGCGCATG----- 2321
QY 768 eGlySerLeuLysLysLeuValGluAsnAlaGluMetLysAsnGlnAsnLeuGluGluAs 788
Db 2322 -AAATCCGCTGAGAGTGTGATGCGGAGGCGAGAGAACAGAAAGAGCCAGCTTCGAGACA 2380
QY 788 pLeuGluThr-----LysThrLysLeuLeuLysGluGlnGluIleG 802
Db 2381 AATGATTCGCTGCGCGAGAAATGCGCCAAAGCTCAAGGCCCGCCGAGCAGCTTCGCGCT 2440
QY 802 nLeuAlaGluLeuArgLysArgAlaAspAsnLeuGlnLysLysValAlaArgAsnPheAsp 822
Db 2441 TAAAGCCGAGAGAAAGCGGCTGAGAGACTCGC----- 2477
QY 822 uSerValSerMetLysAspSerGluLysLeuGluGluIlePheGlnLeuLysGlnSe 842
Db 2478 -----TCCATGTTTCATTTTCAG-----ATGAGACAGCTA----- 2507
QY 842 rLeuSerAspAlaGluAlaValThrArgAspAlaGlnLysGluLysSerPheLeuArgSe 862
Db 2508 -----CGCGAAGGCCACACCCG-----CAGGTGTCGAGCTCCGCGGA 2545
QY 862 rGluAsnLeuGluLysGluLysMetGluAspThrSerAsnTrpLysGlnLysG 882

```

```

Db 2546 CGAAATGCCGCCAAGCAGCAGCAAGAAATGACAGATGAAGAT---GTCCATCAAAAG-- 2600
QY 882 uLysAlaIleSerLeuPheGluLysGlnLeuGluThrGlnLysSerAsnTrpLysLysSe 902
Db 2601 -----CTGCTCTTGGCGACACCAACAGATGACGGCCGCTACGAAAGGTGCGCAGA 2653
QY 902 tGluAlaAspLeuGlnLysGluLeuGlnSerAlaPheAsnGluIleAsnTrpLeuAsnG 922
Db 2654 GGATCCCGAAGATCCACAGAGCTTCAG-----AACAT 2686
QY 922 yLeuLeuAlaGlyLysValProArgAspLeuLeuSerArgValGluLeuGluLysVal 942
Db 2687 CATCTCCAC----- 2696
QY 942 lSerGluPheSerLysGlnLeuGluLysAlaLeuGluGluLysAsnAlaLeuGluAsnG 962
Db 2697 -----AACAGCTCGGAGACCAAGCGGCAAAAGCTCAAGAGGCTTGAGGAC-- 2744
QY 962 uValThrCysLeuSerGluLysPheLeuProAsnGluValGluLysLeuLysAsnG 982
Db 2745 -----ACGGTGGCCAGAGAGTGCAGACCTACACAACTCGCAAACTTTCTGTCAGA 2800
QY 982 nLysSerLysAlaSerGluGluIleMetLeuLeuLysGlnGluGluLysIleHisSerAlaSe 1002
Db 2801 TCTACAGCAACCAATCCGAAGAAATGTCTAAACGAGAGAGCCAGAGAGCGGTGATC 2860
QY 1002 rIleIleSerLysGlnGluIleIleMetGlnGluGlnSerGluGlnIleLeuGlnLeuTh 1022
Db 2861 ACTCGCGCAAGAAACGAAGATTTCTTCTTGAGAACAAAC-----CT 2902
QY 1022 rAspGluValThrHisThrGlnSerLysValGlnGlnThrGluGluGlnLysLeuGlu 1042
Db 2903 CGACACAGTGAC----- 2915
QY 1042 tLysLysMetHisAspAspLeuPheGluLysTrpIleArgAsnLysSerGluAlaGluAs 1062
Db 2916 ---AAGGTGCAACAGCAATG-----GTGGGGCAACCGCCGATCTCGCGTG 2959
QY 1062 pLeuLeuArgGluMetGlu-----AsnLeuLysGlyThrMetGluSerValGluValLys 1081
Db 2960 CGAGTCCGCCAAGCTGAGAAAGCGTCTACGCTGACATGAGCGGCTGAAGAGCTCTGGA 3019
QY 1081 eAlaAspThrLysHisGluLeuGluGluThrIleArgAspLysGluGlnLeuLeuHisG 1101
Db 3020 GACACGCTCAAGAGAGGCGAAGAGCGCAATGCGGATGCAAGCGCTACCAATAGCA 3079
QY 1101 u 1101
Db 3080 G 3080

```

```

RESULT 13
US-09-134-001C-322
Sequence 322, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT FILING DATE: US/09/134,001C
CURRENT APPLICATION NUMBER: 1998-08-13
PRIOR FILING DATE: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 322
LENGTH: 30549
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-322
Alignment Scores:

```

Pred. No.:	3.26e-59	length:	30549
Score:	866.00	Matches:	661
Percent Similarity:	38.24%	Conservative:	697
Best Local Similarity:	18.61%	Mismatches:	1295
Query Match:	5.86%	Indels:	898
DB:	4	Gaps:	157

[illegible]

Db	19819	AAT	-----GATRAAAAAACAATAATTATGTATTAATGCTATTAAAGAAAGCTGACGATTTAAAT	19872
Qy	828	-----	-----ASperSerg1uLyLeuScySglu1u1IerPheGlnLeuLys	840
Db	19873	CATTAATCATCAGCATATGATTAAGTATCATTAATAGACATTACGAAGATTTATTAACAAGATAGAC	19932	841
Qy	841	GlnSerLeuSerAspAlaGluAlaValThrArg	-----AspAlaGln	854
Db	19933	CAGGCGCAATACGAATAATTAATGAGAAATCCAGATTTTAAACAGGCTTTAGACAAATGCTTTA	19992	855
Qy	855	LysGlnCysSerPheLeuArgSerGluAsnLeuGlnLeuLysGluLysMetGluAspThr	874	875
Db	19993	AACGACATATGATAGCTATTAAACAGTCTCAATGTTCACAAAGCCAAACTGTTAAAGGTTAAC	20052	876
Qy	875	SerAsnTrpTyrAsnGlnLysGluLysAlaAlaSerLeuPheGluLysGlnLeuGlu	893	894
Db	20053	ATCAACCATGTGACAACTAGAAAGTTAGCTCAAGAAATTCGAGAAACCAAAAGGCTT	20112	895
Qy	894	-----	-----ThrGluLys	896
Db	20113	AATGATGCTATGAAAGCAATGAGAGATACCATTTATGATTAATCAAGACAAATTCGTAAATAAT	20172	897
Qy	897	SerAsnTyrLysLysMetGluAlaAspLeuGlnLysGluLeuGlnSerAlaPheAsnGlu	916	917
Db	20173	ACCAATTAATACATAATGACAGACTAGCTCAACAAAAATGCTATATATCATGCTAGATTAAT	20232	918
Qy	917	IleAsnTyr	-----	919
Db	20233	ATTAATACATTAATTTGGTACAGACAAATGCGAGAGATGAGTCTTCMAATTAATCAACAGCA	20292	920
Qy	920	-----	-----LeuAsnGlnLeuAlaGluLysValProArgAspLeu	932
Db	20293	ACTCAAGATATTAATATACAGCTTAATATGATTTAAATGAGATCAAAAACTTCAGAGT	20349	933
Qy	933	LeuSerArgValGluLeuGlnLysLysValSerGluPheSer	-----	946
Db	20350	---GCCAAACACAGATGCTTAAACACAAATTAAGTAACTTTACCTTTCGTTTAACTGACACCAAA	20408	947
Qy	947	---LysGlnLeuGlnLysAlaLeuGlnGluLysAsnAlaLeuGlnLysGlnLysValThrCys	965	966
Db	20407	AAACAAAGCTTGGAAACATCTTAAACACAAACAGACAGACAAATTTGCTAAACAG	20466	967
Qy	966	LeuSerGluTyrLysPheLeuProAsnGluValGluCysLeuLysAsnGlnLeuSerLys	985	986
Db	20467	TTAAGTCATGCTAAATTTCTTAATGAGAAATAAGAAATTAAGATTAAAGTTGACAGAGCCAAA	20526	987
Qy	986	AlaSerGluGlnIleMetLeuLeuLysGlnGluGlnIleHisSerAlaSerIleIleSer	1005	988
Db	20527	GGCTTA	-----TTAGTAAACAAATTAAGTAACTAATTAATGAAGATGCTCTT	20574
Qy	1006	LysGlnGlnIleIleMetGln	-----GluGlnSerGlnGlnIleLeuGlnLeuThrAsp	1023
Db	20575	GAAAAAGAAAGCATATGACACAACTATCGAAAAAGGTACAGAAATATTAATTCAGAAAT	20634	1024
Qy	1024	GluValThrHisThrGlnSerLysValGlnGlnIleThrGlnGlnLysGlnIleMetLys	1043	1044
Db	20635	AATCCACAAATAGTAGTACTGATATCATCATGCTACCATTCAGAAATTAATGATGCTGAA	20694	1045
Qy	1044	Lys---MetHisAspAspLeuPheGluTyrTyrIleAsnLys---SerGluAlaGlu	1061	1062
Db	20695	CAAAATCTTCAAGCTGAT	-----AATAAATTTAAGACAAACACACAG	20733
Qy	1062	AspLeuLeuArg	---GluMetCysAsnLeuLysGly	1072
Db	20734	GAAATTTGCAAAAGATGAATCAAAATCTAGACGATTAATTCAGCTCAATTAACAAA	20793	1073
Qy	1073	-----	-----ThrMetGlnSerValGluValLysIleAlaAsp	1083
Db	20794	TTAATTCCAAGATATAGGACAGAACAACTAAACCTGACGATACACAGAACTAGAACGA	20853	1084
Qy	1084	ThrLys	-----HisGluLeuGlnGluThrIleArgAspLysGlnGln	1097
Db	20854	GCAAAGCAATTAACCAAGCTATGCAACACTTTAAACAAAGATATGCGCATAGATGAGTCT	20913	1098

QY 1098 LeuLeuHisglulysLysTyrPhe----- 1105
 Db 20914 ACCTAAATCTGTAAGTCTTCATGATGAGATTCGAGAAAAAGTTAGGCTAGCATAT 20973
 QY 1106 -----PheGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuPro 1123
 Db 20974 GCTGTAGCCAGCTGAGCAACTC-----ATTAATCAACTTAACGACCCCAACTATG 21024
 QY 1124 ProSerLysLeuValGlnGluAsnSerGlnAspProIleGluIleAsnAsp---TyrHis 1142
 Db 21025 GATTAAGTAATATTCAGCTATTCTCAAAAGGTCATTCAGCAAAAGATTCATTGCAC 21084
 QY 1143 AsnLeuIleAlaLeuAlaThrGluArg-----AsnAsnIleMetValCysLeuGluThr 1160
 Db 21085 GGTCGGAATAACTGTCACAAAATCAAGCAGATTCATTAATTAATTAATCAATCACA 21144
 QY 1161 GluArgAsnSerLeuLysGlnGlnValIleAsp----- 1171
 Db 21145 AATTAAATGATTAACAAAAGCAGCATTAATGACTTAATATCATGCTCAAACTTAA 21204
 QY 1172 -----LeuAsnThrGlnLeuGlnSerLeu 1179
 Db 21205 CAGCAAGTGCAGAAATATTCGACAGCTAATTAAGTTAATTAACGAATGGCAGACTA 21264
 QY 1180 GlnAlaGlnSerIleGluLysSerAspLeu---GlnLysProLysGlnAspLeuGlu 1198
 Db 21265 AAAACACTCGTAGAAGACAGCTCAACGTTTCATCAACAAAGTAAATATTAATGAGAT 21324
 QY 1199 GlyLysValLysLeuLeuGluMetGlnLeuLysGlnHis----- 1213
 Db 21325 CCGGAGCTCAAAATATTTATTAATGACTCCATCAAAAAGTCGAGAAATTAATTAACGGC 21384
 QY 1214 LeuThrAspSerGlnLeuSerIleGluLysLeu-----GlnLeuGluAsnLeuGluVal 1231
 Db 21385 ACACAGATGATGTTTAAACAAATTAATGACATGCGATTCCTCAAAATCATTTA 21444
 QY 1232 Thr-----GluLysLeuGlnThrLeuGlnGluMetLysAsn 1244
 Db 21445 ACATAAAACGATTTCATGCTGATCAAAAATTCACAAAACACCAAGATGCAACCAAT 21504
 QY 1245 IleThrIleGluArgAsnGluLeu---GlnThrAsnIleGlnAspLeuLysAlaGluHis 1263
 Db 21505 -----GAAATTAACATTTTAAACAATCTAAACATTTCTCAAAAGACGAGCAT 21555
 QY 1264 AspSerLeu-----LysGlnAspLeuSerGlnAsnIleGlnGlnSerIle 1278
 Db 21556 GATGAGATTAACCTGCTCTCTCAAGAACGAGAGTTTCTATGATTTAAATCATGCTAAA 21615
 QY 1279 GluThrGlnAspGluLeuArgAlaAlaGlnGluGluLeu-----ArgGlnGlnLys 1295
 Db 21616 GCACTTAATGAGCTAGCTCAACTCAAGATGAGATGAGTGTCTTGAACACAGTGTAAA 21675
 QY 1296 GlnLeuValAspSerPheArg----- 1302
 Db 21676 AAATTAAGCGACTTTATCAATGAGATGAAGCGGACAAAATGAAATATGATTAATGCATT 21735
 QY 1303 GlnGlnLeuLeuAspCysSerValGlyIleSerSerProAsnHisAspAlaValAlaAsn 1322
 Db 21736 CAAAAGCTAAAGACATTAACACGGGCTCCAGTAGACGCTTAAGTTAAAGCTACAT 21795
 QY 1323 GlnGlnLysValSerLeuGlnGluValAsnSerLeuGlnSerGluMetLeuArgGlyGlu 1342
 Db 21796 GAAGAT-----GCTTATATGAAATGCAAAATGCTAGA---GAAAGTTTACATGCTAG 21846
 QY 1343 ArgAspGlnGlnIleThrSerCysLysAlaLeuValSerGlnLeuGlnLeuLeuArgAla 1362
 Db 21847 CAA---AACTTCAAGGCTAAATAATCAAGCTGTGCTGAATTAATTAATTAACACA 21903
 QY 1363 HisValLysSerValGlnGluGluAsnLeuGluIleThrLysLysLeuAsnGly----- 1380
 Db 21904 -----TTAAATCTCTGACAG 21918

QY 1381 -----LeuGlnLysGlnIleLeuGluLysSerGlnLeuSerGlnValLeuLysSerMet 1398
 Db 21919 GTTCTTCTGAAAAAACAATAGTTAAACCAAGCATCAACCAACCAAGAGATTCAGAAAGCC 21978
 QY 1399 LeuGlnAsnLeuLysGlnAspAsnLysLysLeuGlnGlnAlaGluIleTyrSerSer 1418
 Db 21979 TTACAAAAGCAAAAGAACTTAATGAAGCTATGAACACATGAAACTGAAATAATTAATA 22038
 QY 1419 LysGlnAsnGlnPheSerLeuGlnGluValPhe-----SerLysSerGlnLys 1434
 Db 22039 AAGAAA---CAATCAAGCTGATAGATGATTAATGTAATCTGACAGCTGCTTCAAGCA 22095
 QY 1435 LeuValAspGlnIleGluValLeuLysAlaGlnLeuLysAlaAlaGlnGluArgGlu 1454
 Db 22096 AATTCAAAATTCGGTTAAATTAATGCTTCAATATTCGCAATCTGCAACCCCAACCAAGAG 22155
 QY 1455 IleLysAspArgAspTyrPheGlnLeuValGlnThrAlaAsnThrAsnLeuValGluGly 1474
 Db 22156 CTTAATTAAGATGTAATTAATTAAGCAACTCAAAAGCTTAAACT----- 22200
 QY 1475 LysLeuGlnThrProLeuGlnAlaAspHisGlnGluAspSerIleAspArgSerGlu 1494
 Db 22201 -----GCTGAAAATATTAATTAATGGGCAATCTTAA 22230
 QY 1495 GluMetGluIleLysValLeuGly---GluLysLeuGlnArgAsnGlnTyrLeuLeuGlu 1513
 Db 22231 TTAGCAGAGGCTAAGTCAGACGGAAATCAAAAGCATGCAACATTTGCAAGATTAACCA 22290
 QY 1514 ArgLeuGlnGluGluLysLeuGlnLeuSerAsnLysLeuGluIleLeuGln----- 1530
 Db 22291 TCACAAAAGATTAACCAACATGATTAATTAATCAAGCTCAAACTAAACCAAGATGAT 22350
 QY 1531 -----LysGluMetGlnThrSerValLeuLeuLysAspArgLeuGln 1544
 Db 22351 GATATCGTAATTAACCTTAACCAATTAAGTACTATG-----ATCAACATCA 22401
 QY 1545 GlnLysLeuGlnSerLeuLeuSerGluAsnIleIleLeuLysGlnAsnIleAspThrThr 1564
 Db 22402 CAATGCTT-----AACATGACAAATACA 22425
 QY 1565 LeuLysHisLysSerAsp-----ThrGlnAlaGlnLeuGlnLysThrGlnGln 1580
 Db 22426 GTAAACAAAATAGTGTTCATTAATGAAGATTCAGCCACCAAGATGCTTAATCAT 22485
 QY 1581 GlnLeuGlnLeuAlaLysAsnLeuAlaIleAlaAspAspAsnCysProIleThrGln 1600
 Db 22486 GCAATTCAGACACCAAAAGATTGATTAACGCT----- 22518
 QY 1601 GluLysGlnThrSerAlaAspCysValHisProLeu-----GluLysLysIleLeu 1617
 Db 22519 -----CATCAACTATCATGATTAATAATCAAAATGAT 22551
 QY 1618 LeuLeuThrGlnGluLeuHisGlnLysThrAsnGlnGlnGlnLysLeuLeuHisGlnLys 1637
 Db 22552 CAAGCTATTGAAAATATCAAAACAGACATTAATGAT-----TTACAGGCTAGT 22599
 QY 1638 AsnGlnLeuGlnAlaIleGlnValGluLeuLysCysGlnValGlnHisLeuMetLysSer 1657
 Db 22600 AATTAACCTATCAACAAGATAA-----AAAGAA 22626
 QY 1658 MetIleGlnSerLysSerSerLeuGlnSerLeuGlnHisGlnLysHisAspThr---Glu 1676
 Db 22627 GCTTCAGACACCACTACAAACCTTAATAGCTTGACCAACGGGCAAAAGATGAGATTTTA 22686
 QY 1677 GlnGlnLeuLeuAlaLeuLysGlnGlnIleGlnValValThrGln-----GluLysLys 1694
 Db 22687 AATCATATTTTCAGTCACACCAACAGAGCAAGTGAAGGAAAAAATTCGAAGTCTTAA 22746
 QY 1695 GluLeuGlnGlnIleThrHisGlnHisLeuThrAlaGlnValAspHisLeuLysGlnAsnIle 1714
 Db 22747 CAATTAATTAATTAATCAATGAAGACACCTTAAGATTAATGATTAATTAATTAATTA 22806
 QY 1715 GluLeuGlnLysAsnPheLysAsnGlnAlaGlnGlnLysThrThrLysGlnGlnCysLeu 1734

Db 22807 CAAAGTACTAGTACTTCAATGAGATCTCTGACACACAAAGCTTATATCAACG--- 22863
Qy 1735 LeuAnGluAnLysGluLeuGluInSerGlnHisArgLeuGluInCysGluLeuGluIn 1754
Db 22864 GTAATTAAGCTAAATAATATATATATGATCAACCAACACAGTAAATGCAATATGATGAG 22923
Qy 1755 LeuMetLysSerLeuLysAspLysGluSerLysLeuGluInThrLeuLysGluSerLysGluIn 1774
Db 22924 ATTCAAAGTCCTTAATGAAAGTTAAACAACTAAAGATATTTA---CAGGTATCA 22960
Qy 1775 LysValIleAsnLeuAnLysGluInMetGluMetValMetLeuGluMetGluInLys 1794
Db 22981 AAAGTCGTAACGACACAGATGCTCAGACATTAATATGCTTAATATCTTAAT 23040
Qy 1795 AsnSerGlnArgThrValIle-----AlaGluArgAspGluLeuGluInAspAspLeu 1811
Db 23041 CAAGGCAAAAGGTAACTTGGAACTAAAGTTCAAACTTAATCTTAATCTAGACCAAGTA 23100
Qy 1812 ArgGluSerValGluMetSerIleGluThrGlnAspAspLeuArgLysAlaGluInLys 1831
Db 23101 CAAAAGTAGTTCATTAAGCAATCACTTAATGATGCGAGTGAATAATTTAGATGATGCT 23160
Qy 1832 Leu-----GluGluInLysAspLysValGluGluInLeuThrSerGln 1845
Db 23161 TTAAGTGTATGACGCAATTAACAAACGAGTATATATTAATGAAGATCTCTCA 23220
Qy 1846 -----IleSerValLeuGluGluLysLysSer 1854
Db 23221 CAAGTAACTTGTAGTATACAGATAGAGTAAACATTAAGTGTGCAACCAAACT 23280
Qy 1855 LeuLeuGluAnGluMetLeuThrValIleAspLysGluInThrLeuSerGlu--- 1873
Db 23281 -----CCAAATATGCTCCAACTAATATTAACACTATGCTGATGATAATTAATCTAAGCT 23334
Qy 1874 ArgAspAspLeuAsn-----GlnSerLysGlnHisLeuPheSerGlu 1887
Db 23335 AAAAAGCTTTTACATGCGCTCAAAACCTAAACAGCTCAACACAGCTCCATCACT 23394
Qy 1888 IleGluThrLeuSerLeuSerLeuLysGluLysGluPheAlaLeuGluInLys 1907
Db 23395 ATTAATCAATGACT-----GGCTTAACCAAGCTCAAAA 23430
Qy 1908 AspLysAlaAspAlaAlaArgLysThrIleAspIleThrGluLysIleSerAsnIleGlu 1927
Db 23431 GAACAATTAAT-----CAAGAAATCAACAAACTCAACCCGTTCTGAAGTA--- 23478
Qy 1928 GluInLeuLeuGluGlnAlaThrAsnLeuLysGluThrLeuThrGluArgGluSerLeu 1947
Db 23479 CATCAAGTAATTAATTAAGCAACAGCTTAATGATCAATGAATACCTTTACGTCAGT 23538
Qy 1948 IleGlnCysLysGluGlnLeuAlaLeuAnThrGlnHisLeuArgGluThrLeuLysSer 1967
Db 23539 ATTACTGATGACATGAAGTTAAACAAACAGATTAATCAATCAATGAAGATGTTGCTAAT 23588
Qy 1968 Lys-----AspLeuAlaLeuGlnLysMetGluGlnLysArgAspGluAlaAlaAsn 1984
Db 23599 CAAGTCGATTAATCAACATGCGCTGATGCTGAATAAACAATAATCAACAAACATCTAAT 23658
Qy 1985 LysValIleAlaLeuThrGluLysMetSerSerLeuGlu---GluGlnIleAsnGluAn 2003
Db 23659 CCAACT-----ATGATCTTATAGAGTGGAACGTCGAACATCAAT 23700
Qy 2004 ValThrThrLeuLysGluGly---GluGluLysGluInThrPheThrLeuGluInArgPro 2022
Db 23701 GTAAAAATTTTAAAGATGCACTTCATGCTGACAGTGA---TTGAATGACATTAATAAT 23757
Qy 2023 SerLysGlnInSerSerSerGlnMetGluLysLeuArgGluSerLeuLysThrLysAsp 2042
Db 23758 TCAAAAACCTTTTGCAGTCATCACTATGATTAACCTCAATCAAGCT----- 23802
Qy 2043 LeuGluGluGluGluAlaGluLysGluLysSerGluAlaThr-----AsnGluIleLys 2060
Db 23803 ---CAAAAAGACATTAATCATGAAATTGACAAAGCAACATATGTTTCAACATTAAT 23859
Qy 2061 AsnLeuThrAlaLysIleSerSerLeuGluGluIleLeuGlnAsnAlaSerIleLeu 2080
Db 23860 AATATCTATTAACAAAGCGCAACGCTTAATTAATGATGATGAAAAA-----CTT 23907
Qy 2081 AsnGluAlaValSerGluAnLysGluAnLeuArgHisLysSerLysGluInLeuValSerGlu 2100
Db 23908 AAAGATATCGTTGCTCAACAGATTAATGAGACATCAACAAATTAATTAACAGAGAT 23967
Qy 2101 LeuGluInLeuSerLeuThrLeuLysSerArgAspHisAla----- 2114
Db 23968 AGTACACCTCAAAATATGATGACAAAGATTAATCATGACCAATCAATCATGATCA 24027
Qy 2115 PheAlaGlnSerLysArgGluLysAspGluAlaValAsnLysIleAlaSerLeuAlaGlu 2134
Db 24028 GTAGCAAAACCTGACATGCTCATGACGAATAGAAATGACATTAATTAACATTAAGCAT 24087
Qy 2135 GluIleLysIleLeuThrLysGluMetAspGluPheArgAspSerLysGlu----- 2151
Db 24088 GCCATCATGTCATCGATGAGAA---CATTAATTAACAAACGCAAAAGAAATGCAAAAC 24144
Qy 2152 -----SerLeuGluInLysInSerHisLeuSerGluGluLeuCysThrTy 2167
Db 24145 TTATGATTAATTAATGATTAACGATTAATGACCAACAAAGAGATGCATTAATGATG 24204
Qy 2168 LysThrGluLeuGluMetLeuLysGluGlnLysGluAspIleAsnAn-----LysLeu 2185
Db 24205 GTTAATGAAGCTCAACAAAGAAAGAGTGAACCACTTCAAAAGTCTCAACCTTTA 24264
Qy 2186 AlaGluLysValLysGluValAspGluLeuLeuGlnHisLeuSerSerLeuLysGluIn 2205
Db 24265 AATGACGCTATGACACATTTAAGAAACAGCATTCAAATCATCATCCGTAGACAAAG 24324
Qy 2206 LeuAspGluIle-----GlnMetGluLeuArgAsnGluLysLeuArgAsn 2220
Db 24325 AGCAATATATTAATTAATGCAAGTATGCTTAATAAAGCAATTAATGACGCACTTGA 24384
Qy 2221 TyrGlu-----LeuGluGluLysMetAspIleMetGluLysGluIleSerValLeuArg 2238
Db 24385 GTCGAAATATTAATCAATGACACACATCCACATTGGATTAAGAAATA-----ATTAG 24438
Qy 2239 LeuMetGlnAsnGluProGluGlnLysGluAspValAlaGluArgMetAspIleLeu 2258
Db 24439 CAACTTAAGGATGCTGAATCAACGAAATATGACTTA---AATGCGGTTGAATTAATA 24495
Qy 2259 GluSerArgAsnGluGluIleGlnLysLeuMetGluLysIleSerAlaValTy 2278
Db 24496 GATGCTGATTAACCAACGACATCAATCGATACCTATGATGCTCAATTAATCAACGA 24555
Qy 2279 GlnHisThrLeuLeuSerSerLeuSerGluGluInLysGluThrGluAlaHis--- 2297
Db 24556 CAAACAAAGCATTAAATGAATAATTAATTAACGAGTTACCAAGAACTGAACTGCGGCT 24615
Qy 2298 -----LysHisCysMetLeuAsnIleLysGluSerLeu--- 2308
Db 24616 ATTATGGCAAGCAAACTACTGATCATGATGATGAGAAATTAAGAAAGATATCAAA 24675
Qy 2309 -----SerSerThrLeuSerArgSerPheGluSerLeuInThr 2321
Db 24676 GATAAAGACAGTCAAAACAGTCAAGTAACTATTAATGAAGATTCGATGCTCAACAA 24735
Qy 2322 GluHisValLysLeuAnThrGlnLeuGlnThrLeuLeuAsnLysPheLysValTy 2341
Db 24736 ACATTCGATTAACGCGCTGATGATGATGACAGAAATCTTAAT-----CAACAGTAAAT 24789
Qy 2342 ArgThrAlaAlaValLysGlu---AspHisSerLeu-----IleLysAspTy 2358
Db 24790 CCAACTTATCATATGAGATTAAGATGATGATGATCAACGAAGTAAATCAACGCAAA 24849
Qy 2359 AspLeuAlaAlaGluGlnLysArgHisAspGluLeuArgLeu-----GluLeuGln 2375
Db 24850 CAAGTCAGAGTAAACAAAACTTATCAAACTATGATGATTAAGTAAAGAAATTAAGT 24909

QY 2376 CysLeuGluGlnHisGlyArgLysTrpSerAspSerAlaSerGluLeuLeuPheCys 2395
 Db 24910 AAATGGATGATTAAATCACAACAAAGACGTTCAATATCAATTAATTAATCTGCT 24969
 QY 2396 Gln-----Ile 2397
 Db 24970 AAAACGACAAACAGAGTCCCAAGCAATGAAAGCAAAATCAATTAATCATGCAATG 25029
 QY 2398 GlnPheLeuAsnGluLeuLeuPheLysAlaAsnIleGlnSer-----Val 2414
 Db 25030 AAAGCACTTAACAAATATATAAATGCAAGATTAAGTTTATGATAGTATGCAATTCAT 25089
 QY 2415 GlnAspAspPheSerGluValGlnValPhe-----LeuAsnGlnValGlySer--- 2430
 Db 25090 AACGAAGATCAACCTGAAAAAAGGCGGTATCAACAAAGCTATTAATATCATGTTATCATTC 25149
 QY 2431 -----ThrLeuGlnGluLeuGlnHisLys 2439
 Db 25150 ATTCAATAGCAACAACAAATCTGAAATGCATCCACACAGATACATATAGCATATCATGAT 25209
 QY 2440 LysGlyPheMetGlnTrpLeuGlnLupPheGlyAspLeuHisValAlaLysLysLeu 2459
 Db 25210 CTCGAACACAGCTCAA-----ATAACTTACATGGTGAT---CAGAACTT 25251
 QY 2460 SerGluGlyMetGlnGlnLupAsnArgArgLysLeuThrIleGlnLeuLeuThrLys 2479
 Db 25252 GCTCATGCAACAACAGATGCCGCTAATGTAATTAATGCTCAATTCATCTT----- 25302
 QY 2480 ArgLeuAlaValAlaGlnSerLysIleGlnArgGluIleThrValTyLeuAsnGln 2499
 Db 25303 -----AATGTGCTCAAGCTGAGTAATGATTAATATACGAAATCA 25341
 QY 2500 PheGluAlaLysLeuGlnGlnLysLysGluLupAsnLysGluLeuMetArgArgMetGlu 2519
 Db 25342 ---AATGCTCAACACGCGGAAGTTGCCAAAGACTTAAATGCTCAAGCTCTTGAT 25398
 QY 2520 HisHisGlyProSerAlaSerValMetGluGlnLupAsnAlaArgLeuLeuGlyIleLeu 2539
 Db 25399 AAAGCTATGAAACACATCAACAGTAGTCTCATATAAATATATATTAATGAAACGATAGT 25458
 QY 2540 LysThrValGlnAspGluSerLysLysLeuGlnSerArgIleLysMetLeu----- 2556
 Db 25459 AAAATTTTAAATGAAGATTCAAAATATCAACACATACGATGATTAATGCTGATGCC 25518
 QY 2557 GlnAsnGluLeuAsnLeuValLysAspAlaMetHisLysGlyLysValAlaIle 2576
 Db 25519 GAACAACTACTTATATCAGACAAACAAATCCAACTTA---GAACTTATTAAGTCGATATT 25575
 QY 2577 LeuGlnAspLysLeuLeuSerArgAsnAlaGlnLupLeuAsnAlaMetGlnValLys 2596
 Db 25576 GTTAAGATATATGCTTAGCT---AACGAAAAAATACTATTGCGCGAAGAAAACHTCA 25632
 QY 2597 LeuThrLysLysGlnAspAsnLeuGlnAlaAla---MetLysGluLeuGlnAsnLeuGln 2615
 Db 25633 TATGACAAATCAATGCAAAATGATGAATTAACATATGATATCTTATATATGACAA 25692
 QY 2616 LysMetValAlaLysGlyAlaValProtyr-----LysGluGluIleAspAsnLeu 2632
 Db 25693 AACCAATCTATAAAGATATGATTTCTCACGACGACATTAAGCACTAAGATTAAACAATT 25752
 QY 2633 -----LysThrLysVal 2636
 Db 25753 CTGCAACAAGCTAAATCTTGTATGAGGACGATGAATCACTTGAAGATTAACACTCAAGTA 25812
 QY 2637 ValLysIleGluMetGluLysIleLysTySerLysAlaThr----- 2650
 Db 25813 GTGATTACAGATCTACTTGGCTAATTACACGAAAGCTTTCAGAGATTAAGAAAAAGAAAA 25872
 QY 2651 ---AspGlnGluIleAlaTyLeuLysSerCysLeuGlnAspLysGluGluGly----- 2667
 Db 25873 GTAGACCAAACTGATATCATGCTCAAGGATTAAT---GATTAATAATATGCTCAAT 25929

QY 2668 -----LeuArgArgLeuLysGluLeuArgArgAla-----GlnAla 2680
 Db 25930 GTAAGTTAGATCAAGTACGACAGACACTAGCAAAATTAATCAAGATCAGAAAACCTC 25989
 QY 2681 AspAsnAspThrThrValCysValProLysAspArgGlnLysAlaSerThrProVal 2700
 Db 25990 GATGCTGATCAGCCGAGTGAAGAGCTAAAGTTACTGATCTTAATCAACAA----- 26037
 QY 2701 ThrCysGlyGlySerGlyIleValGlnSerThrAlaMetLeuValLeuGlnSerGlu 2720
 Db 26038 -----ATTGATCAATTAACATCATCTTATTCATTAACAAACA 26076
 QY 2721 LysAlaAlaLeuGluArgGluLeuSerHisTyLysLysLysTyLysHisLysSerArg 2740
 Db 26077 ACTCG-----AAGAAAGTGTAAAAACGCAACAAACCTAGAAAGAAATCGCTACT 26127
 QY 2741 ThrMetSerSerGluAspArgLysThrLysAlaLysSerAspAlaHisSerSer 2760
 Db 26128 GTTAGTAAACAATGCTCAGGATTAACAAAGTAAATGAGTAAATTAACAATTCATTAAT 26187
 QY 2761 HisThrGlySerSerHisArgGlySerProHisLys----- 2772
 Db 26188 CATGCTGATTCCTGTGAAAAATAGTATTAATTAATAGCAACGACGACGCAAAATCATC 26247
 QY 2773 -----ThrGluThrTyArgHisGlyProValThrProGluArgSerGluMetProSer 2790
 Db 26248 GCTTATGATGACAGACACTGAACATGACAA----- 26277
 QY 2791 LeuHisLeuGlySerProLysLysSerGluSerSerThrLysArgValAlaSerProAsn 2810
 Db 26278 ---GATATACAAAAAATCAACGCAACCAAAATGAACAAACAAACGCTTA----- 26325
 QY 2811 ArgSerGluIleTySerGlnLeuValMetSerProGlyLysThrGlyMetHis----- 2828
 Db 26326 CAACATTAATATATATGACAAACATGCTTAATAGTTTGAAGATTAATATGATGCTAGA 26385
 QY 2829 -----LysHisIleLeuSerProSerLysValGlyLeuHisLysLysArgAla 2844
 Db 26386 CCAAGAGCTTTAGAAATATATCAAAATCACTAGAAATAAACAAGCTCAAAAGCTGCT 26445
 QY 2845 LeuSerProAsnArgSerGlu-----MetProThrGlnHisValIleSerPro 2860
 Db 26446 TTAGAGATTAAGTAAAGCAAAATGCGATGATTAATTAATGAATTAAGAAATATGCTCAACAG 26505
 QY 2861 GlyLysThrGlyLeuHisLysAsnLeuThrGluSerThrLeuPheAspAsnLeuSer 2880
 Db 26506 GGC----- 26508
 QY 2881 ProCysLysGlnGlnLysValGlnLupAsnLeuAsnSerProLysGlyLysLeuPheAsp 2900
 Db 26509 -----ACAAACCTCAATGACATTAATGCGTAATGCTAAC 26544
 QY 2901 ValLysSerLysSerMetProTyCysProSerGlnPhe-----Phe 2914
 Db 26545 GCAATCGTTAATTAAC-----TATGCTCAACCAACAAAGCAAGTAAATTAATTAACGCC 26598
 QY 2915 AspAsnSerLysLeuGlyAspPheSerGluLeu-----AsnThrAlaGlnLupSerAspAsn 2932
 Db 26599 GATTAACCTACGCAAAAGATTAATCTTACTCAACCTATCAACATATGACGAGTACCTCAAC 26658
 QY 2933 LysSerGlnAlaGluAsnTrpTrpLysGluAla 2943
 Db 26659 AAAACTCAAGCTCAAGACTTGAATTTCAATGCA 26691

RESULT 14
 US-09-166-350-27
 : Sequence 27, Application US/09166350A
 : Patent No. 6440663
 : GENERAL INFORMATION:
 : APPLICANT: Scanlan, Matthew
 : APPLICANT: Chen, Yao
 : APPLICANT: Stockert, Elisabeth
 : APPLICANT: Old, Lloyd


```

Db 3335 -----GTCAACTTCAAAAGAAAAGAAAGAGCTTCAGAAAACCATGCAGAAATTA 3385
Oy 2060 LysAsnLeuThrAlaLysIleSerSerLeuGluGluIleLeuGlnAsnAlaSerIle 2079
Db 3386 GAG-----CTGGTTAAAGAAAGATGGCCCAACAAACCATG 3421
Oy 2080 LeuSnglAlaValSerGluArgGluAsnLeuArgHisSerLysGlnGluValSer 2099
Db 3422 ATGAATATGCAAAATAGCTGATTATGAA-----CGTTGATGAAA 3460
Oy 2100 GluLeuGluGluLeuSerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLys 2119
Db 3461 GAACTTAATATCA-----AAGTTAACTAATAATAAAC----- 3490
Oy 2120 ArgGluLysAspGluAlaValAsnLysIleAlaSerLeuAlaGluIleLysIleLeu 2139
Db 3491 -----AACAGATAGAGATTTGGAGCAAGAAATATAAAATTCAA 3529
Oy 2140 ThrLysGluMetAspGluPheArgAspSerLysGluSerLeuGlnGlnSerSerHis 2159
Db 3530 -----AAACAGAAACAGAAACCCCTACAGAAAGAAATTAATCTCA 3568
Oy 2160 LeuSerGluGluLeuLeuCysThrLysThrGluLeuGlnMetLeuLysGlnGlnLysGlu 2179
Db 3569 TTACAGCTTCACATCAACATATGAGAAAGAAAACACCAAAATCAACATGCTGTGTG 3628
Oy 2180 AspIleAsnAsnLysLeuAlaGlu---LysValLysGluValAspGluLeuLeuGlnHis 2198
Db 3629 AAAACCAAAAGAAAGAACTGGAGATTCACAAACAGCAAGCAAACTGATCACTTAATCTTCA 3688
Oy 2199 LeuSerSerLeuLysGluGlnLeuAspGlnIleGlnMetGluLeuArgAsnGluLysLeu 2218
Db 3689 ---GCATCTTTAAAGGTGAGCTGGAGCAAGCCAGCAGCAAGTGAAGTCTAATAATA 3745
Oy 2219 ArgAsnThrGluLeu---CysGluLysMetAspIleMetGluLysGluIleSerValLeu 2237
Db 3746 CAGCTGGCTGAATATACATCAGAGAAAGCAAAATCCAGAGCAC----- 3790
Oy 2238 ArgLeuMetGlnAsnGluProGlnGlnGluLysAspValAlaGluArgMetAspIle 2257
Db 3791 -----CTGAAGAACCTTCGCGAAGACACACAG-----CGTAGC 3823
Oy 2258 LeuGluSerArgAsnGlnGluIleGlnGluLeuMetGluLysIleSerAlaValLysSer 2277
Db 3824 CTAGTGCTATCCAGCAGAGAGTGACAGACTACAGAAAGAGTCCGTGCTCCAAAGCA 3883
Oy 2278 GluGlnHisThrLeuLeuSerSerLeuSerSerGluLeuGlnLysGluThrGluAlaHis 2297
Db 3884 GAACAGAGCTACTGTAACTCT-----GAATTCGAGAGCTAC 3919
Oy 2298 LysHisCysMetLeuAsnIleLysGluSerLeuSerThrLeuSerArgSerPheGly 2317
Db 3920 AAA-----GTCCGAGTTCAATATGTTCTAAACCAACAAAGAAATAATCTATGCTCT 3970
Oy 2318 SerLeuGlnThrGlnHisValLysLeuAsnThrGln---LeuGlnThrLeuLeuAsnLys 2336
Db 3971 CAGGCTGAAGCTGAGGCGCTTAACAGAAAGGAGACATCTGGAAATCTGATTATACAG 4030
Oy 2337 PheLysValValLysArgThrAlaAlaValLysGluAspHisSerLeuIleLysAspThr 2356
Db 4031 CTAAATAATC-----AAATTCACAGATACCAAAATTAATCAAGATTAT 4075
Oy 2357 GluLysAspLeuAlaAlaGluGlnLysArgHisAspGluLeuArgLeuGlnLeuGlnCys 2376
Db 4076 GTATCTGAACCTCAACATTCAGCTGACATCTGATACATG----- 4117
Oy 2377 LeuGluGlnHisGlyArgLysTrpSerAspSerAlaSerGluLeu----- 2392
Db 4118 CTAAAGAGGCAACAGATCTCTCAGAGAAACTGTCTCAAGAGGCCGAAGCTCCGGGAA 4177
Oy 2393 LysPheCysGluIleGluPheLeuAsnGluLeuLeuPheLysLysAlaIleIleGln 2412

```

```

Db 4178 AAATGTGTCATATACAG-----TCAGAGAACTGATGATGAATCTGACATACAG 4231
Oy 2413 SerValGlnAspAspPheSerGluValGlnValPheLeuAsn----- 2426
Db 4232 ACTGTGATCAGCTTAACATCCAGAGAGGCTCTTCAAAATAGCTTCCGAGATCAAGTG 4291
Oy 2427 -----GlnValGlySerThrLeuGlnGluLeuGlnHis 2438
Db 4292 CGACATTTGCAGAGAAACACAGAAAGACATGGAGACTTTCACAGACGCTCTCCAG 4351
Oy 2439 LysLysGlyPheMetGlnThrLeuGln-----PheGlyAspLeu 2452
Db 4352 ATGAGACAGACGCTCTCCAGCTTAAGATGAACGACACAGAACGCCAGTTCCTCT 4411
Oy 2449 ----- 2448
Db 4412 CAACATTTTGAAGAACCTTCGAGAAAGAGAAACAGACGCTCCCTTCAACATG 4471
Oy 2453 HisValAspAlaLysLeuSerGluGlyMetGlnGlnLysArgArgIleAlaSer 2472
Db 4472 CACACTGTAAACCCGGGAGAGGAGAGAGACATGAGACATGATACGAGTCTGTCT 4531
Oy 2473 ThrIleGlnLeuLeuThrLysArgLeuLysAlaValGlnSerLysIleGlnArgGlu 2492
Db 4532 TCCGCGACGACATACACAGACTTTAGAGCAGTG----- 4567
Oy 2493 IleThrValThrLeuAsnGlnPheGluAlaLysLeuGln-----GluLysLys 2508
Db 4568 -----CTTAAGCTCTCCGAACTAACTGAGACCTTCATATGAGCATCTGAA 4615
Oy 2509 GluGlnAsnLysGluLeuMetArgArgMetGlnHisHisGlyProSerAlaSerValMet 2528
Db 4616 TTACCAAGAAAGAAATGTGTTAGAGCTCAAGCTTCACACAAAGAGTGCAT----- 4669
Oy 2529 GluGluGluAsnAlaArgLeuGluLysIleLeuLysThrValGlnAspGluSerLysLys 2548
Db 4670 -----CACTTAAACGCGCTGCTCGGGAAGCAAGAAACCAATCAAT 4714
Oy 2549 LeuGlnSerArgIleLysMetLeuGluAsnGluLeu---AsnLeuValLysAspAspAla 2567
Db 4715 CTATAGAGCAATTAATGCTTCAAAAGTGAAATAGAAAGATGGAAGATCAAGAG 4774
Oy 2568 MetHisLysGlyGlu-----LysValAlaIleLeuGlnAspLysLeu 2581
Db 4775 CGAGAGAAAGTGCAGCTAACTGGAATACTGGAAGACGCTTCTGCGATGCTTTTC 4834
Oy 2582 LeuSerArgAsnAlaGluAlaGlu-----LeuAsnAlaMetGln-ValLys 2596
Db 4835 TTGAACCAAGCTAGTGAAGAAAGAGAGACTTCTCTGTTATGAATACGATCTTGACAGCTC 4894
Oy 2596 IleuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLysGluIleGluAsnLeuGlnLys 2616
Db 4895 AGCCCTGAAGAAAGGAAACCTTCTCGGCTGCTCAAGGTGAGAGAAAGAAATGCTTCC 4954
Oy 2616 smetValAlaLysGlyAlaValAlaProTyrLysGluLysLeuAsp----- 2630
Db 4955 CGTCTTCTGATGGCATCTCATCTTCACTAGTGTGCTGAGCTCGATAGTGTATGTA 5014
Oy 2631 -----AsnLeuLysThrLysValValLysIleGluMetGln 2642
Db 5015 AGGAATATTTTATTAACCAATAGATCTATTTCAAAAAGGTTCAGCTATATTACCA 5074
Oy 2642 ulysIleLysTyrSerLysAlaThrAspGlnGluLeuAlaTyrLeuLysSerCysLeu 2661
Db 5075 CAAT-----TCTTTGCAAAAAGTGTGA 5099

```

RESULT 15
 US-09-724-517-1
 ; Sequence 1, Application US/09724517
 ; Patent No. 6379941
 ; GENERAL INFORMATION:
 ; APPLICANT: Berand, Christophe
 ; APPLICANT: Freedman, Richard

; TITLE OF INVENTION: No. 6379941el motor proteins and methods for
 ; FILE OF INVENTION: their use
 ; FILE REFERENCE: 1031
 ; CURRENT APPLICATION NUMBER: US/09/724,517
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US/09/641,807
 ; PRIOR FILING DATE: 2000-08-17
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3837
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1226)...(1337)
 ; OTHER INFORMATION: n = a, t, c, or g
 US-09-724-517-1

Alignment Scores:

Pred. No.: 1,48e-58 Length: 3837
 Score: 841.50 Matches: 342
 Percent Similarity: 42.40% Conservative: 244
 Best Local Similarity: 24.75% Mismatches: 537
 Query Match: 5.70% Indels: 259
 DB: Gaps: 47

US-09-150-867-1 (1-2954) x US-09-724-517-1 (1-3837)

QY 7 VallyValCysValArgValArgProLeuIleGlnArgGluGlnGlyAspGlnAlaAsn 26
 DB 19 GlnAAAGTCTGTAAAGAAATTAAGACTCTGTCGCAAGAGCTCTTCAATATCATCAA 78
 QY 27 LeuGlnIlePheValGlnGlnAsnThrIleSerGlnValAspGly---ThrLysSerPhe 45
 DB 79 GTTGTGTGAGATTAATCCAAAGCAGCAGCACTTACTTGGAGAGATAGAGTCTTC 138
 QY 46 AsnPhesprValPheAsnSerHisGlnSerThrSerGlnIleGlyGlnIleAla 65
 DB 139 ACTTTGATTTGTTGTTGTCGCAAAATTCACATGATGAGTTTAAACATGATATA 198
 QY 66 ValProIleIleArgSerAlaLeuGlnGlyTyrAsnGlyThrIlePheAlaTyrGln 85
 DB 199 AAGCCCTAGTGTGTACTCATCATGAGGCTATATGCAACTGTGTTGCTATGACAA 238
 QY 86 ThrSerSerGlyThrThrThrMetGly-----ThrProAsn 99
 DB 259 ACTGATCTGGGAAGACATACCATTTGAGGGGCCATTTGCTTCAAGTTGGAGGCG 318
 QY 100 SerLeuGlyIleIleProGlnAlaIleGlnValPheLysIleIleGlnIlePro 119
 DB 319 CAAAGGGTTCATCTCCAGCATTCAGAAATATTTCAAGCACTCTGAAACATCCT 378
 QY 120 AsnArgGluPheLeuValArgValSerThrMetGluIleTyrAsnGluThrValLysAsp 139
 DB 379 AGCATTTGACTTATGTAAGATCTTATATGAGAGTATGACAGGAAGCCTTAAGGAT 438
 QY 140 LeuLeuCysAspArgArgLysLysProLeuGluIleArgGlnAspPheAsnArgAsn 159
 DB 439 CTCTTGAATTTGAGACATCATGAGATCTTCACATTCGAGAGAAAGTGAAGAAAC 498
 QY 160 ValTyrValAlaAspLeuThrGlnGluLeuValMetValProGlnHisValIleGlnTyr 179
 DB 499 ACAGTATTTGTTGGGCGCAAGATGCGCATGTGAGAGTGCAGGTGAAGATGAGTCTT 558
 QY 180 IleLysLysGlyLysAsnArgHisTyrGlyGluThrLysMetAsnPhisSerSer 199
 DB 559 TTGGATGGGAGATGAGCAGACATACAGGTACCACTCAATATGATGAGCAGCTCAGC 618
 QY 200 ArgSerHisThrIlePheArgMetIleValGlnSerArgAspArgAsnProThrAsn 219
 DB 619 AGATCAATGCAATTTTACATCAGCATTTGTCAAAGTTCATATAAATATGAGAGCCT 678

QY 220 SerGlnAsnCysAspGlyAla-----ValMetValSerHisLeuAsnLeu 234
 DB 679 GAA-----GATGATCATGTGATATCCCGGACATATGTGTCGCAAGTTCACATT 729
 QY 235 ValAspLeuAlaGlySerGlnArgAlaSerGlnThrGlyAlaGluValArgLeuLys 254
 DB 730 GTGATTTGGCAGATCAGAAAGTAAGTAACCAACGGGGAATCTGCTGAAGGTTCCAA 789
 QY 255 GlnGlyCysAsnIleAsnArgSerPheIleLeuGlyGlnValIleLysLysLeuSer 274
 DB 790 GAATTCATTTCAATCAATATGATGATTCCTGCTTGAAGAAATGTAATACCGCTCTGG 849
 QY 275 Asp---GlyGlnAlaGlyPheIleAsnTyrArgAspSerLysLeuThrArgIleLeu 293
 DB 850 GACCCAGCAGGAAGATTCACATATTCATATGAGGATCTAAATACCGGCTCTG 909
 QY 294 GlnAsnSerLeuGlyGlnAlaValThrValIleLeuLysThrIleThrProValSer 313
 DB 910 AAAGATTTCTGGAGAGCTGTAAGACTGTATGATGATGATGATGATGATGATGATGAT 969
 QY 314 -----PheAspGluThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHisValArg 331
 DB 970 TCGAATTTGATGATGCTTAAATTTCTCAATATGCGCAAGAGCAGCAGCAATTAGA 1029
 QY 332 AsnThrProHisValAsnGluValLeuAspAspGlnAlaLeuLeuLysArgTyrArgLys 351
 DB 1030 AACAAACCCCATGTAACCTTCACGCCCCGAGCTCAGACCGT---ATAGATGAATGAATTT 1086
 QY 352 GlnIleLeuAspLeuLysLysGlnLeuGlu----- 361
 DB 1087 GAGATTAATTTGCTTCAGAGCTTTGCAAGAGCAGCAGCAGCTGCTGACCCAAATACC 1146
 QY 362 -----AsnLeuGluSerSerSerGluThrLysAlaGlnIleAlaLysGlnGlnHis 379
 DB 1147 CAGATCAATGAGAGAGGATCCTGATACAAATAGAGTATCTCTGTAGAGCAAGTA 1206
 QY 380 ThrGlnLeuAlaGluIleLysGlnLeuHisLysGlnArgGluAspArgIleThrPhe 399
 DB 1207 GCTCAGCTTCACAGAGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1266
 QY 400 LeuThrAsnIleValAlaIleSerSerGlnGln----- 412
 DB 1267 NNN 1326
 QY 413 -----GlnAspGlnArgValLysArgLysArgValThr----- 424
 DB 1327 NNN 1386
 QY 425 -----ThrAlaProGlyLysIleGlnAsnSerLeuHis 435
 DB 1387 TTTGATACATTTGGGCGATTATATTATATCCATCAAGA---CAAGATTCAGAGAG 1443
 QY 436 AlaSerGlyValSerAspPheAspMetLeuSerArgLeuProGlyAsnPhisSerLys 455
 DB 1444 GTCCACACAAAGTCGCTATGTAATCTGTCGATCGAATATTTCTGATTCGAACACGA 1503
 QY 456 AlaLysPheSerAspMetProSerPheProGluIleAspAspSerValCysThrGluPhe 475
 DB 1504 AGTCAGATGCTG---TTGGTCAATAGAGACAGATAGAGTCTCCATCCCAATTT 1560
 QY 476 SerAspPheAspAlaLeuSerMetMetSerSerHisGlyIleAspAlaGluTyrPhe 495
 DB 1561 TCTGATTAACAGTATGAT-----GAAGAATCCAGAAAGCCCAAGAAATCTGGAAC 1611
 QY 496 LeuAlaSerLysValThrHisArgGluLysThrSerLeuHisGlnSerMetIleAspPhe 515
 DB 1612 AGATGTAAGAGCTTCATGATGATTCAGAAACCAAGCTCTGTTGTTCCCTGTGAATTTG 1671
 QY 516 GlyGlnIleSerAspSerValGlnPheHisAsp----- 526
 DB 1672 AGTGAATCTCAGATGAAACACAAAGTCAATTTGGAGAAATGAAGATTTGAATTTGAT 1731
 QY 527 -----SerSerLysGlnAsnGlnLeuGlnTyrIleuProLysAspSerGlyAspMet 543

Db 1732 TGTCACGAGAGTCAGAAATTTGCAAAATTAAGAAATTCAGAACGATCACT 1791
 QY 544 AlAGlucysArgLysAlaSerPheGluLysGluIleThrSerLeuGlnGlnLeuGln 563
 Db 1792 ACTGAGCTGAA-----CAAAATGAGAGAACTTCAATTAACATCAAG 1836
 QY 564 SerLysGluGluLysLysGluLeuValGlnSer----- 575
 Db 1837 ATGAAGGAGAGTCGATTAAGAAATTAATAAACAAGTAATGCCAAGCTCTTAAGC 1896
 QY 576 -----PheGluLeuLysIleAlaGluLeuGlu-----GluGlnLeuSerValLys 590
 Db 1897 AAGCAGATCTCTTGAAGTAACAAGCTAGCATGATGCAAGACGACGAAAGCTGAA 1956
 QY 591 -----AlaLysAsnLeuGlnLysValThrAsnSer-----ArgGluHisSer 604
 Db 1957 CTGATGGAACACAAACACCTACAGAGCTGGAACAAAGATCTTCTGATGTTCA 2016
 QY 605 IleAsnAlaGluValGlnThrAspValGluLysGluValAlaArgLysGluMetSerVal 624
 Db 2017 ATGAAGGTAATTAACAGAAAGAGTTGTTAAAGATGATGCTGCAAGCTGAGAGTT 2076
 QY 625 LeuGluAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSerSerValAspGly 644
 Db 2077 CAGGCTCTGACAGAAACACCAACAGATAGTAAGAACTGGCATCCTGTCATCCAAAT 2136
 QY 645 LysArgLysSerSerSerHisAspGluLysIleGluHisArgLys-----Met 660
 Db 2137 GAGAACCTGCTGAATGACCTGAGCAGAGTGATGATCACAATGAAATATCAAAAGATACG 2196
 QY 661 LeuGluGlnLysIleValAspLeuGluLysPheIleGluAsnLeuAsnLysLysSerGlu 680
 Db 2197 CTACAAGAAATACACGAAAGAAATGAAAGAAAGCAACGATGATGACGATATTAG 2256
 QY 681 AsnAspLysGluLysSerSerGluGlnAspPheMetGluSerIleGlnLeu----- 697
 Db 2257 CGGAGCCGCAAAATCAAGAA-----ATCAATTAATAAACAAGCA 2298
 QY 698 CysGluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeuMetArg 716
 Db 2299 CAGAGAAAGGCTTAAACCGAAGCTGAGACCTGATGATGATGATGATGAAAGGAGA 2358
 QY 717 -----AspAsnPheAspAsnIleIleLeuGluAsnGluLysThrLeuLysArgLys 732
 Db 2359 AAAGTTCCTTGAAGATATGACCATCTCCAGAAATTGCAATGCAAAAGAAAGGTTA 2418
 QY 733 IleAlaAspLeuGluArgSerLeuLysGluAsnGlnGluIleThrAsnGluPheGlu 750
 Db 2419 GATGAGAGAGTAGAGAAAGTTCTGAACCAACGCAAGAAATTAAGAGACCTGGAACAGAC 2478
 QY 751 IleLeuGluLysGluThrGlnLysGluHisGluAlaGlnIleHisGluIleLysSer 770
 Db 2479 TTAAAGAAACGAGGAGCCATGATTCTTAAGAAAGAGCTGCTGTTACAGAGAAAGATCAC 2538
 QY 771 LeuLys--LysLeuValGluAsnAlaGluMetLysTyrAsnGlnAsn----- 784
 Db 2539 CTGAAATATAGAAATGAGATGATGATGAGCCCTTAACACAGATATGTTGAAATATCA 2598
 QY 785 -----LeuGluGluAspLeu----- 789
 Db 2599 ACTGCGCTGAACCTTACTGGAACAGAGTTGCTGAAAGAAATGTCACAGCTCCAGACAGT 2658
 QY 790 -----GluThrLysThrLysLeuLeuLysGluGlnGluIleGlnLeuAlaGluLeuArg 807
 Db 2659 ACAGCTGAGGAGAAACAAAGATTTCAACACAGTTGAAGTC-----CTCAG 2706
 QY 808 LysArgAlaAspAsnLeuGlnLysLysValAlaArgAsnPheAspLeuSerValSerMetGly 827
 Db 2707 AAAGAAAGATCACTCCAAAGAGCAGACAGATGTGATGAAAGAAATTAAGATGT 2766
 QY 828 AspSerGluLysLeuGluGlu-----IlePheGlnLeuLysGlnSerLeuSerAsp 845

Db 2767 AGAGTGTATACCTGCAAGAAACATGTTCTTCCAACTGGAAGAGGATAGAGCT 2826
 QY 846 AlAGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerLysAsnLeu 865
 Db 2827 TTGGAAGCTGCATTTGAA-----TACAGGAATGAAGATATC 2862
 QY 866 GluLeuLysGluLysMetGluAspPheSerAsnTyrAsnGlnLysGluLysAlaAla 885
 Db 2863 CAGATTCGCCACAAAGCTACCTTAGACATCA---TTCATTAACCTCTCTGCTGACAGCA 2919
 QY 886 SerLeuPheGluLys-----GlnLeuGluThrGluLysSerAsn 898
 Db 2920 AATGCTTGGAAGAACTACCTGCTGACCTGCTGTTGAGATTAAGAACTTCTTTTCA 2979
 QY 899 TyrLysLysMetGluAlaAspLeuGlnLysGluLeuGlnSerAlaPheAsnGluLeuAsn 918
 Db 2980 TAT-----TTCATTAAGGTGTGTG 2997
 QY 919 TyrLeuAsnGluLeuAlaGluLysValProArgAspLeuSerArgValGluLeu 938
 Db 2998 AATTGGCA-----GAGCT 3012
 QY 939 GluLysLysValSerGluPheSerLysGlnLeuGlu--LysAlaLeuGluLysAsn 957
 Db 3013 GAACGGAACACACAGTTATATATGAAGAAATGAAGAAATGTTCTGGAACGGGATAT 3072
 QY 958 AlaLeuGluAsnGluValThrCysLeuSerGluTyrLysPheLeuProAsnGluValGlu 977
 Db 3073 ATGTTCTGGAATTAAGATCTGACCTGACATCTAATA-----TTGCA 3117
 QY 978 CysLeuLysAsnGlnIleSerLysAlaSerGluGluIleMetLeuLysGlnGluLys 997
 Db 3118 TGT-----GACCGAGAGCTGACCTCCAGCAAAAG 3147
 QY 998 GluHisSerAlaSerIleLeuSerLysGlnGluIleMetGlnGluGlnSerGluGln 1017
 Db 3148 GAAACAGAA-----CAAAAGATGCGAGTTGCTATTCATCATTCATCAAAACAA 3195
 QY 1018 IleLeuGlnLeuThrAspGluValThrHisThr--GlnSerLysValGlnGlnThrGlu 1036
 Db 3196 GATGAGAGAGGCAATTAATGGAACCTTCAAAACATTAATGAAGATTAATCCAGAGTTGGA 3255
 QY 1037 GluGlnTyrLeuGluMetLysLysMetHisAspAsp-----Leu 1049
 Db 3256 AAAGATCTTATTTCTATTAAGAAACACCGCGGATCATTAAGAAACTTAAGAACTG 3315
 QY 1050 PheGluLysTyrIleArgAsnLys-----SerGluAlaGluAspLeu 1063
 Db 3316 GTAGCGGAAGCAATTCGGCGGCAACCTAGCACCATCAGATATCAAGAGCTGAGATGGA 3375
 QY 1064 LeuArgGluMetGluAsnLeuLysGluThrMetGluSerValGluValLysIleAlaAsp 1083
 Db 3376 GTCTGAAAGCCAGAGGA-----GAGAGCATGCTTTCAGAAATTAATAATGGCATCC 3429
 QY 1084 ThrLysIleGluLeuGluGluThrIleArgAspLysGlu-----GlnLeu 1098
 Db 3430 AGACCTGAAGATGAATTAAGTGAAGAGAAAGAAAGAAAGAGTTCACAGACAGC 3489
 QY 1099 LeuHisGluLysLysTyrPhePheGlnAlaMetGlnThrIlePheProIleThrProLeu 1118
 Db 3490 TTAAACACACAGCAAAATCTCTAAAGCTGCGGAGATATGCCAATTAACCTCAATTT 3549
 QY 1119 SerAspSerLeu--ProPheSerLysLeuValGluGluLysSerGluAspProIleGlu 1137
 Db 3550 CATAGCTTTTATGACACCCCGAGTGGCATATAGTTAGTATGAGAAAT----- 3597
 QY 1138 IleAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCys 1157
 Db 3598 -----AAACAGAAACAGATATACAGTTTCA 3627
 QY 1158 LeuGluThrGluArgAsnSerLeuLysGluGlnValIleAspLeuAsnThrGlnLeuGln 1177
 Db 3628 AAATCTCACAGCTGACCTGATCCCAATTCAGTTGTGGGAATATGTGAGACGACTTCAT 3687

```
QY 1178 SerLeuGlnAlaInserIleGluYserAspLeuGlnLysEroLYsGlnAspLeuGln 1197
      :::: :::: :::: ::::
Db 3688 GGTCTCACACCTGTAAACTGTGTGCGAAAAGATTAAGTCAATTC----- 3735
QY 1198 GluGlyGluValLysLeuLeuGluMetGluLeuLysGlyHisLeuThrAspSer 1217
      :::: :::: :::: ::::
Db 3736 -----GCCTTGGAACTATTCATTCGACGCTCCAGCTCT----- 3768
QY 1218 GluLeuSerIleGluYsLeuGlnLeuGluAsnLeuGluValThrGluYsLeuGlnThr 1237
      :::: :::: :::: ::::
Db 3769 GGACTTGGCATGTGATCAATGGCTGCTGATTCATCGAAGATATCTAGAAACCAAGGAC 3828
QY 1238 LeuGln 1239
      ::::
Db 3829 TTAATA 3834
```

Search completed: May 23, 2003, 15:16:20
Job time : 824 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 23, 2003, 15:03:25 ; Search time 709 Seconds
(without alignments)
5501.616 Million cell updates/sec

Title: US-09-150-867-1

Perfect score: 14769
Sequence: 1 MSECDAKVCVRRVRLQRE.....QAEWWYEAKKETAPCKTS 2954

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame_plus_p2n.model -DEV-xlp
-O=/cgn2_1/USPTO.spool/US09150867/rnph_21052003_165914_21023/app_query.fasta.1.3143
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=trnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPOST=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09150867_ECGN_1_1.518_6rnph_21052003_165914_21023
-DSPU=6 -ICPU=3 -NO_XLPT -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-NDBLOCK=0 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10
-XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	3615.5	24.5	7992 9	US-09-893-519A-140 Sequence 140, App
2	3613.5	24.5	8493 9	US-10-071-766-51 Sequence 51, Appl
3	1153	7.8	10300 10	US-09-960-253-145 Sequence 145, App
4	1152	7.8	10096 10	US-09-960-253-163 Sequence 163, App

5	1139	7.7	10190 10	US-09-864-864-292 Sequence 292, App
6	1139	7.7	10211 10	US-09-954-456-1153 Sequence 1153, App
7	1139	7.7	10211 10	US-09-967-768A-186 Sequence 186, App
8	1024.5	6.9	12313 9	US-10-171-311-7 Sequence 7, Appl
9	1024.5	6.9	12313 9	US-10-171-311-3 Sequence 3, Appl
10	1022.5	6.9	12337 9	US-10-171-311-5 Sequence 5, Appl
11	1022.5	6.9	12462 9	US-10-171-311-1 Sequence 1, Appl
12	998	6.8	11677 9	US-10-082-830-134 Sequence 134, App
13	900.5	6.1	6386 9	US-10-098-841-40 Sequence 40, Appl
14	891.5	6.0	7497 10	US-09-960-253-175 Sequence 175, App
15	891.5	6.0	7792 12	US-10-044-090-359 Sequence 359, App
16	888.5	6.0	3366 9	US-09-938-842A-2651 Sequence 2651, App
17	848	5.7	6773 10	US-09-864-864-336 Sequence 336, App
18	836	5.7	14800 10	US-09-954-456-1601 Sequence 1601, App
19	829	5.6	7065 9	US-09-991-466-115 Sequence 115, App
20	829	5.6	7065 10	US-09-874-923-115 Sequence 115, App
21	797	5.4	15231 10	US-09-917-800A-1505 Sequence 1505, App
22	795	5.4	10625 9	US-10-023-219-3 Sequence 3, Appl
23	795	5.4	10625 10	US-09-727-384-5 Sequence 5, Appl
24	786.5	5.3	6586 10	US-09-954-456-1115 Sequence 1115, App
25	786.5	5.3	6586 10	US-09-954-456-1790 Sequence 1790, App
26	776	5.3	8948 10	US-09-735-105-119 Sequence 119, App
27	776	5.3	8948 10	US-09-850-716A-119 Sequence 119, App
28	776	5.3	9588 10	US-09-897-778-119 Sequence 119, App
29	776	5.3	9588 10	US-09-954-456-1848 Sequence 1848, App
30	762	5.2	7596 10	US-09-954-456-2215 Sequence 2215, App
31	754	5.1	13957 10	US-09-782-378A-22 Sequence 22, Appl
32	754	5.1	13957 10	US-09-880-107-2284 Sequence 2284, App
33	721.5	4.9	6861 9	US-10-171-311-161 Sequence 161, App
34	721	4.9	6900 9	US-10-171-311-163 Sequence 163, App
35	721	4.9	8930 10	US-09-954-456-1805 Sequence 1805, App
36	720.5	4.9	9274 10	US-09-885-535-3 Sequence 3, Appl
37	719.5	4.9	5937 9	US-09-927-597-3 Sequence 3, Appl
38	718	4.9	4780 10	US-09-962-436-287 Sequence 287, App
39	718	4.9	18846 10	US-09-815-242-8898 Sequence 8898, App
40	708.5	4.8	5835 9	US-09-927-597-1 Sequence 1, Appl
41	702	4.8	5848 12	US-10-044-090-299 Sequence 299, App
42	699.5	4.7	4858 10	US-09-954-456-733 Sequence 733, App
43	695.5	4.7	10302 10	US-09-782-378A-23 Sequence 23, Appl
44	691.5	4.7	9477 10	US-09-815-242-8513 Sequence 813, App
45	687.5	4.7	31096 7	US-08-781-968A-59 Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-893-519A-140
; Sequence 140, Application US/09893519A
; Publication No. US20030027243A1
GENERAL INFORMATION:
APPLICANT: ANADYS PHARMACEUTICALS, INC.
APPLICANT: THOMPSON, Craig
APPLICANT: MOORE, Jeffrey
APPLICANT: BUDDMAN, Ed T.
APPLICANT: BRADLEY, John
APPLICANT: DESILVA, Thamara
APPLICANT: HARRIS, Sandra
APPLICANT: KOMARNITSKY, Svetlana
APPLICANT: MENDILLO, Marc
APPLICANT: MOORE, Daniel
APPLICANT: MCCOY, Melissa
APPLICANT: SANDERSON, Karen
APPLICANT: HAO, Tarig
APPLICANT: ZHU, Shuhao
APPLICANT: LONG, Fan
APPLICANT: DAVIDOV, Eugene
TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
FILE REFERENCE: 0342/16548-US2
CURRENT APPLICATION NUMBER: US/09/893,519A
PRIOR APPLICATION NUMBER: US 60/215,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457

PRIOR FILING DATE: 2000-08-10
 NUMBER OF SEQ ID NOS: 146
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 140
 LENGTH: 7992
 TYPE: DNA
 ORGANISM: Homo sapiens
 PUBLICATION INFORMATION:
 DATABASE ACCESSION NUMBER: Human Genbank/Z15005.1
 DATABASE ENTRY DATE: 1993-01-10
 RELEVANT RESIDUES: (1)..(7992)
 US-09-893-519A-140

Alignment Scores:
 Pred. No.: 2,26e-223 Length: 7992
 Score: 3615.50 Matches: 996
 Percent Similarity: 50.34% Conservative: 572
 Best Local Similarity: 31.97% Mismatches: 932
 Query Match: 24.48% Indels: 615
 Gaps: 86

US-09-150-867-1 (1-2954) x US-09-893-519A-140 (1-7992)

QY 1 MetSerGluGlyAspAlaValCysValArgValArgProLeuIleGlnArgGlu 20
 Db 1 ATGGCGAGGAGGAGGAGCGCGCTGCGTGGAGTGGCGGCGTGAACAGCAGAGAA 60
 QY 21 Gln-----GlyAspGlnAlaAsnLeuGlnTrpLysAlaGlyAsnAsnThrIleSerGln 38
 Db 61 GAATCACTTGGAGAAATGCGCAAGTTACTGAAAACCTGACAAATATGTCATTATCA 120
 QY 39 ValAspGlyThrLysSerPheAsnPheAspArgValPheAsnSerHisGluSerThrSer 58
 Db 121 GTTATGAGGAGTAATCTTCAATTTGATCGTGTCTTATGATGTAATGAACACCAA 180
 QY 59 GlnIleTrpGlnIleAlaValProIleLeuSerAlaLeuGlnGlyTrpArgGly 78
 Db 181 AATGTGATGAGAAATAGCAGCCCAATCATGCTGCAATCAAGCTCAATAGT 240
 QY 79 ThrIlePheAlaArgGlyGlnThrSerSerGlyLysThrTrpMetMetGlyThrPro 98
 Db 241 ACATATATTCCTATGACAGACGCTTCAGAAAACATATACCATGATGGGTTCGAA 300
 QY 99 AsnSerLeuGlyIleIleProGlnAlaIleGlnGluValPheLysIleIleGlnIle 118
 Db 301 GATCATTTGGAGATTAACCAAGGCAATTCATGCAATTTCCAAAAATTAAGAACTT 360
 QY 119 ProAsnArgGluPheLeuLeuArgValSerTrpMetGluIleTrpAsnGluThrValLys 138
 Db 361 CCGATAGGGAATTTCTTACGATGATCTTACATGGAATATATACATGAACCATTACA 420
 QY 139 AsnLeuLeuGlyAspArgArgGlyLysProLeuGluIleArgGluAspPheAsnArg 158
 Db 421 GATTTACTCTGTGCACTCAAAAATGMAACCTTATATTTCAGAGAAAGTGCATAGG 480
 QY 159 AsnValTrpValAlaAspLeuThrGlnGluLeuValMetAlaProGluHisValIleGln 178
 Db 481 AATGTGATTTGCTGATCTCACAGAGAGATGTATATACATCAGAAATGCTTTGAAA 540
 QY 179 TrpIleLysGlyGlyLysAsnArgHisTrpGlyGluThrLysMetAsnPheHisSer 198
 Db 541 TGGATTACAAAGGAGAAAGACAGCATTAATGGAAGAAACAAATATGATCAAGAGC 600
 QY 199 SerArgSerHisThrIlePheArgMetIleValGluSerArgAspArgAsnProThr 218
 Db 601 AGTCGTCATACCATCTTTAGATGATTTGGAAGCAGAGAGAGGCTGAACCTTCC 660
 QY 219 AsnSerGluAsnGlyAlaValMetValSerHisLeuAsnLeuValAspLeuAla 238
 Db 661 -----AATTGCAAGGATCTGTTAAGTATCCATTTGAATTTGTTGATCTTGCA 711
 QY 239 GlySerGluArgAlaSerGlnThrGlyAlaGluGlyValArgLeuLysGluGlyCysAsn 258
 Db -----

Db 712 GCGAGTGAAGAGCTGCTCAACAGCGGCTGAGGTGTGCGGCTCAAGAGGCTGTAAT 771
 QY 259 ILAsnArgSerLeuPheIleLeuGlyGlnValIleLysLysLeuSerAspGlyGlnAla 278
 Db 772 ATTAATCAAGCTTATTTATTTGGACAAAGTATCAAGAACTTGTGTGAGCAAGTT 831
 QY 279 GlyGlyPheIleAsnTrpArgAspSerLysLeuThrArgIleLeuGlnAsnSerLeuGly 298
 Db 832 GGGGTTCATTAATATATATGAGATAGCAAGTAAACAGAAATCTTCAGATTCCTGGGA 891
 QY 299 GlyAsnAlaLysThrValIleIleCysThrIleThrProValSerPheAspGluThrLeu 318
 Db 892 GGAATCCAAAGACACGTATATCTGCACAAATATCTCCATATCTTTGTATGAACCTCT 951
 QY 319 SerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGlu 338
 Db 952 ACTGCTCCAGTTTGGCACTACTGCTAAATATATGAGAAATCTCTTATGTTAATGAG 1011
 QY 339 ValLeuAspAspGluAlaLeuLeuLysArgTrpArgLysGluIleLeuAspLeuLysLys 358
 Db 1012 GTATCAACATGATGAGCTCTCTGAAAAGGTATAGAAAAGAAATATATGATCTTAAGAAA 1071
 QY 359 GlnLeuGlnAsnLeuGlnLysSerSerGluThrLysAlaGlnAlaMetAlaLysGluGlu 378
 Db 1072 CAATTAGAG-----GAGGTTCTTTAGAGACCGCGGCTCAGCAATGGAAGAACCA 1125
 QY 379 HisThrGlnLeuLeuAlaGluIleLysGlnLeuHisLysGlnGluArgLysArgIleTrp 398
 Db 1126 TTGGCCCACTTTTGGAGAGAAAGATTTGCTTCAGAAAGTACAGAAATGGAATTTGAA 1185
 QY 399 HisLeuThrAsnIleValAlaLysSerSerGlnGluSer---GlnGlnAspGlnArgVal 417
 Db 1186 AACTTAAACAGGATGCTGGAGCCTCTCTCCCTCAGCTGACACAGAAATTAAGGCT 1245
 QY 418 LysArgLysArgArgValThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAlaSer 437
 Db 1246 AAAAGAAAGAGAGATGTTCTGTGCTTGGCAAAATTAACAAATTAAGAACTCAAAAC 1305
 QY 438 GlyValSerAspPheAspMetLeuSerArgLeuProGlyAsnPheSerLysLysAlaLys 457
 Db 1306 TATGCAAGATCAATTTAAT-----ATACCAACAAATGTATACCAACAAACAT 1353
 QY 458 PheSerAspMetProSerPheProGluIleAspAspSerValCysThrGluPheSerAsp 477
 Db 1354 AACCTTTCTATAAATTTATTAACGAAATTAAGAAATCTGTCTGTCACAGCTGATGTT 1413
 QY 478 PheAspAspAlaLeuSerMetLeuAspSerAsnGlyIleAspAlaGluTrpAsnLeuAla 497
 Db 1414 TTCTAGTAACACTCTTATATCATTAAGT-----GAGATGAAATGGAATTCAGCA 1461
 QY 498 SerLysValThrHisArgGluLysThrSerLeuHisGlnSerMetIleAspPheGlyGln 517
 Db 1462 ACMAAGCTACTAATATCAGAG----- 1482
 QY 518 IleSerAspSerValGlnPheHisAspSerSerLysGluAsnGlnLeuGlnTrpLeuPro 537
 Db 1482 ----- 1482
 QY 538 LysAspSerGlyAspMetAlaGluCysArgLysAlaSerPheGluLysGluIleThrSer 557
 Db 1482 ----- 1482
 QY 558 LeuGlnGlnGlnLeuGlnSerLysGluGluGluLysGluLeuValGlnSerPheGlu 577
 Db 1482 ----- 1482
 QY 578 LeuLysIleAlaGluLeuGlnGluGlnLeuSerValLysAlaLysAsnLeuGluMetVal 597
 Db 1483 -----AATATGAAAGGTGAG 1497
 QY 598 ThrAsnSerArgGluHisSerIleAsnAlaGluValGlnThrAspValGluLysGluVal 617
 Db 1498 TTGAACCTCA----- 1506

QY 5318 ValArgLysGluMetSerValLeuClYaspSerGlyTyrAsnAlaSerAsnSerAspLeu 637
 Db 1506 ----- 1506
 QY 638 GluAspSerSerValAspGlyLysArgLeuSerSerHisAspGluCysIleGluHis 657
 Db 1506 ----- 1506
 QY 658 ArgLysMetLeuGluGluHisIleValAspLeuGluGluPheIleGluAsnLeuAsnLys 677
 Db 1506 ----- 1506
 QY 678 LysSerGluAsnAspLysGluLysSerSerGluGluAspPheMetGluSerIleGluLeu 697
 Db 1506 ----- 1506
 QY 698 CysGluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeuMetArgAsp 717
 Db 1507 -----CTTCGTGCT 1515
 QY 718 AsnPheAspAsnIleIleLeuGluAsnGluThrLeuLysArgGluIleAlaAspLeuGlu 737
 Db 1516 GACTATGATTAATCTGGATTAGACTATGACAACTACGACAAAGAAAAGAAATGGA 1575
 QY 738 ArgSerLeuLysGluAsnGluGluThrAsnGluPheGluIleLeuGluLysGluThrGln 757
 Db 1576 TTGAATTAATAAGAAAAGAAATGATTGGATGATTTGGCTCTAGAAAAGAAACTAA 1635
 QY 758 LysGluHisGluAlaGluLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn 777
 Db 1636 AAAATTAAGAGATGCAACCTAATTCATGAACTTGAATTAAGAAATTTTGTATAGCAT 1695
 QY 778 AlaGluMetTyrAsnGluAsnLeuGluGluAspLeuGluThrLysThrLysLeuLys 797
 Db 1696 CGAGAGATATTAATCAAGATCTTGAGATGACTGCTGAAAGTAGAGCTGCTTGA 1755
 QY 798 GluGluGluIleGluLeuAlaGluLeuArgLysArgAlaAspAsnLeuGluLysVal 817
 Db 1756 GAAAAGAGACCGAGATTAGAGACTACAGAAATACATAGACTCTCAAAAGCTGAAAT 1815
 QY 818 ArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysLeuGluIlePhe 837
 Db 1816 ATAAAAATGGACTGTCTACTACTCTG-----GAACCATTTGAGACCCAAA 1863
 QY 838 GluLeuLysGluSerLeuSerAspAlaGluAlaValThrArgAspAlaGluLysGluCys 857
 Db 1864 CAAATGAAAGCAGACTCTGTTGATGCTGAACCTGATGCCCTTGATGCCAAGAGAAATCA 1923
 QY 858 SerPheLeuArgSerGluAsnLeuGluLeuLysGluLysMetGluAspThrSerAsnTrp 877
 Db 1924 GCCTTCTTGAAGTGAATAATCTGAGTGAAGAGAAAAGAAAGAACTTGCAACTCA 1983
 QY 878 TyrAsnGluLysGluLysAlaIleSerLeuPheGluLysGluLeuGluThrGluLysSer 897
 Db 1984 TACAAAGCAAAATGCAATGATATTCACCTATATCAAAAGCCAAATTTGGAGGCAAAA----- 2037
 QY 898 AsnTyrLysLysMetGluAlaAspLeuGluLysGluLeuGluGluSerAlaPheAsnGluIle 917
 Db 2038 -----AAGAAATGCAAGTTGATCTGAGAAAATAATTACATATGCTTTAAAGAGATA 2091
 QY 918 AsnTyrLysAsnGluLeuLeuAlaGluLysValProArgAspLeuSerArgValGlu 937
 Db 2092 ACAAAACTCACTCCCTTATATAGAGGCAAGTTCCAAAAGTTTGCTCTGTAATTTGGAA 2151
 QY 938 LeuGluLysLysValSerGluPheSerLysGluLeuGluLysAlaLeuGluGluLysAsn 957
 Db 2152 TTGGAGCAAGAAAGATTACTGATCTTCACAGAAAGAACTAATAAAGAAATTTGAAGAAATGAA 2211
 QY 958 AlaLeuGluAsnGluValThrCysLeuSerGluTyrTrpPheLeuProAsnGluValGlu 977
 Db 2212 GCTTGGGGGAGAGATCAATTTTGTCTTCAGAAATTTGAATCTTACCTTCTGGAAGTGA 2271

QY 978 CysLeuLysAsnGluIleSerLysAlaSerGluGluIleMetLeuLeuGluGluGlu 997
 Db 2272 AGCTGAGGAAAAGATATCAAGACAAATCTGAGAGCTCCATATTAATCACTGCAAAA 2331
 QY 998 GluHisSerAlaSerIleIleSerLysGluGluIleIleMetGluGluGluSerGluGlu 1017
 Db 2332 GATAAATGTTTCTGAGATGATTCATTAAGAGAGATAGAGTTCAACAGTTTACTTGAA 2391
 QY 1018 IleLeuGluLeuThrAspGluValThrHisThrGlnSerLysValGluGluThrGluGlu 1037
 Db 2392 ATTGGAACCAAAAGATGACCTACCAACTACACAGTGAATTAATTAAGCACTGATCA 2451
 QY 1038 GluTyrLeuGluMetLysLysMetHisAspAspLeuPheGluLysTyr----- 1053
 Db 2452 GAATTCAAAATTTCAAAACCTTCATATGACCTTGACCAAAAGATTAACATGCTCTT 2511
 QY 1054 -----IleArgAsnLysSerGluAlaGluAspLeuArgGluMetGluAsnLeu 1070
 Db 2512 GAGGAGATGAGAGAAATGATTCAGAGAAATAGTTAATCTCTGTAAGAAAGCCAAAATTT 2571
 QY 1071 LysGluThrMetGluSerValGluValLysIleAlaAspThrLysHisGluLeuGluGlu 1090
 Db 2572 GATTCGAGTTGGGTGCTTTGAAGCCGAGCTTCTTCAAGACCCAAAGAACTTCAGAG 2631
 QY 1091 ThrIleArgAspLysGluGluLeuLeuHisGluLysLysTyrPhePheGluAlaMetGlu 1110
 Db 2632 AAACACCTGAGCTTCAGAAAAGACTAATGATGAGAGAACGCTGACAAACATTAAGA 2691
 QY 1111 ThrIlePheProIleThrProLeuSerAspSerLeuProSerLysLeuValGluGlu 1130
 Db 2692 AAT-----AGAGATTCCTCCCTGCAAACTGAGAAAGGAGAAACACTGATTTACTGAG 2745
 QY 1131 AsnSerGluAspProIleGluIleAsnAspTyrHisAsnIleAlaLeuAlaThrGlu 1150
 Db 2746 AAACGCGCAAACTTTGAA-----GAGTAAAAACTTTTAACTCAGAAAGAAAGATAT 2799
 QY 1151 ArgAsnAsnIleMetValCysLeuGluGluThrGluArgAsnSerLeuLysGluGluVal 1169
 Db 2800 CTAATAACAACTCCAAAGAAAGCTTGCAATTTGAGAGGAGCAACAAAGATATTTGAC 2859
 QY 1170 -----IleAspLeuAsnThrGluLeuGlu-----SerLeuGluAlaGlu 1182
 Db 2860 GATACGTGTAATCATGATATATGATACCTACAGAACATTAATCAAAATGCTTGAGCTCTG 2919
 QY 1183 SerIleGluLysSerAspLeuGluLysProLysGluAspLeuGluGluGluValLys 1202
 Db 2920 AAACAACTCAAGAAACATTAATACACTTAAATCGAAAATTTCTGAG--GAAGTTTCC 2976
 QY 1203 LeuLeuLeuGluMetGluLeuLeuLysGluLysLeuThrAsp-----SerGlu 1218
 Db 2977 AGGAATTTGCATATGAGAGAAATATACAGAGAAACTAAGATGAATTCACGCAAAAGATG 3036
 QY 1219 LeuSerIleGluLysLeuGluLeuGluAsnLeuGluValThrGluLysLeuGluThrLeu 1238
 Db 3037 GTTGCAATGATTAATAAAACAG-----GATTTGGA--GCTAATAAATACCCAACACTA 3087
 QY 1239 GluGluGluMetLysAsnIleThrIle----- 1247
 Db 3088 ACTGCAGATGTTAAGGATTAATGAGATTAATGAGCAACAAAGAAAGATTTTCTTTAATA 3147
 QY 1248 --GluArgAsnGluLeuGluThrAsnPheGluAspLeuLysAlaGluHisAspSerLeu 1266
 Db 3148 CAGAGAAATAATGAATCCAAACAATGTTAGAGAGCTGTTATAGCAGAAAGAAACAATTTG 3207
 QY 1267 LysGluAspLeuSerGluAsnIleGluGluSerIleGluThrGlnAspGluLeuArgAla 1286
 Db 3208 AAGACTGACCTTAAGGAATAATTTGAATGACCAATTTGAAAACAGGAAGATTAAGACTT 3267
 QY 1287 AlaGluGluLeuArgGluGluGluLysGluLeuValAspSerPheArgGluGluLeu 1306
 Db 3268 CTTGGGAGTGAACCTTAATAAGCAACAGATGATCTGCACAAAGAAAGCAATGCCATA 3327
 QY 1307 AspCysSerValGlyLysSerSerProAsnHisAspAlaValAlaAsnGluGluVal 1326

QY	2300	CysMetLeuAsnIleLysGlnSerLeuSerSerThrLeuSerArgSerPheGlySerLeu	2319
Db	6390	-----	6390
QY	2320	GlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPheLysVal	2339
Db	6391	-----	6411
QY	2340	Val-----TyrArgThrAlaAlaValLysGlnSerPheHisSerLeuLysAspTyrGlu	2357
Db	6412	CGTTAAAGCTAAGTTCACAAAATTAACAACAACAACACATGATGCATCAATTAATTGGAA	6471
QY	2358	LysAspLeuAlaAlaGlnGlnLysArgHisAspGluLeuArgLeuGlnLeuGlnCysLeu	2377
Db	6472	ATGCAATTTTATTGATGAAAGTGGAAAAGCAAAAAGCAATTCGTAATTTAAATACACACCTT	6531
QY	2378	GlnGlnHisGlyArgLysTrpSerAspSerAlaSerGlnGluLeuLysPheCysGluIle	2397
Db	6532	CACACAGAT-----TGATGTACCATCCACAGATTAAGGATGTCMAATTG	6579
QY	2398	GluPheLeuAsnGlnLeuLeuPheLysLysAlaAsnIleIleGlnSerValGlnAsp	2417
Db	6580	AACCAAGATATGATCTA-----CATATTGAGGAATTCCTAAGAT	6621
QY	2418	PheSerGluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGluLeuGlu	2437
Db	6622	TTCCTCAGAAAGTGAG-----TTCCTCAGCATTAAGACTGAATTTCAACAAGTACTAAGT	6675
QY	2438	HisLysLysGlyPheMetGlnTrpLeuGlnGlnPheGlyAspLeuHisValAspAlaLys	2457
Db	6676	AATAGCAAGAATAAGACACAGTTTTTGGAAAGTGCTTAATTAATCTGTTTGAATAGAA	6735
QY	2458	LysLeuSerGlnGlyMetGlnGlnGlnLysAsnArgIleAlaSerThrIleGlnLeu	2477
Db	6736	AAGCTTAATAAATGGCATCGATCGAAGAAGAAATGATAGATTTTCAAGTCAATTTCTTT	6795
QY	2478	ThrLysArgLeuLysAlaValValGlnSerLysIle-----GlnArgGluIleThr	2494
Db	6796	AATAACACAAATTAATGCCATATGATGATCAACACAGAGTTTGAGAAAGCAAGTCAACC	6855
QY	2495	ValTyrLeuAsnGlnPheGlnAlaLysLeuGlnGlnLysLysGlnGlnAsnLysGlnLeu	2514
Db	6856	ATA---TCCAAAGAGTGGGAACAGCACTGAAATTCAGTAAGAAGAAATGAAAACTA	6912
QY	2515	MetArgArgMetGlnHisIleGlyProSerAlaSerValMetGlnGlnLysAsnAlaArg	2534
Db	6913	TTTAAAACTACCAAACTGAAGACTCTCTGGCATCTGTGCCACAGTTAAT-----	6966
QY	2535	LeuLeuGlyIleLeuLysThrValGlnAspLys-----SerLysLys	2548
Db	6967	-----CCTACACACACACATAGAATTCCTCATGTTACATCAAGAGCT	7011
QY	2549	LeuGln-----SerArgIleLysMetLeuGlnAsnGlnLysAsnLeuValLysAsp	2565
Db	7012	ACACAGTTAACCACAGACAGAAATTCGACAGCTGGAAAAATTCACGTCATGAAGCTAAAGAA	7071
QY	2566	AspAlaMetHisLysGlyLysValAlaIleLeuGlnAspLysLeu--LeuSerArg	2584
Db	7072	AGTGTATGCATTAAGGAAGCAAGATTTAAGAATGAGGAAGCAACTTGGAGTGACTAAT	7131
QY	2585	AsnAlaGluAlaGlnLeuAsnAlaMetGlnValLysLeuThrLysLysGlnAspAsnLeu	2604
Db	7132	GACATTAATGCAAACTT-----CAACCCAAAGTTCATGATCAATTAATAGCCCTT	7182
QY	2605	GlnAlaAlaMetLysGluIleGlnAsnLeuGlnLysMetValAlaLysGlyAlaValPro	2624
Db	7183	GAATAAAACAAAAGAGACATTCAGAGTACTTCAGACAAGAGTGGCTTTAGGAGCTAAAGCA	7242
QY	2625	TyrLysGlnGluIleLeuAspAsnLeuLysThrLysValValLysIleGlnMetGlnLysIle	2644
Db	7243	TATTAAGAAGAAATTTGAAGATCTCAAAATGAAGCTTTGTGAATAAGACCTTAAGCAAAATG	7302
QY	2645	LysTyrSerLysAlaThrAspGlnGlnIleLeuLysLeuLysSerCysLeuGlnAspLys	2664

```

:      NUMBER OF SEQ ID NOS: 144
:      SOFTWARE: PERL Program
:      SEQ ID NO 51
:      LENGTH: 8493
:      TYPE: DNA
:      ORGANISM: Homo sapiens
:      FEATURE:
:      NAME/KEY: misc_feature
:      OTHER INFORMATION: incyte ID No. US20020192678A1 441283.3
US-10-071-766-51

Alignment Scores:
Pred. No.:      3,26e-223
Score:          3613.50
Percent Similarity: 50.34%
Best local Similarity: 31.94%
Query Match:    24,47%
DB:             9
               86

```

```

US-09-150-867-1 (1-2954) x US-10-071-766-51 (1-8493)
QY 1 MetSerGIuIAspAlaValLysValCysValArgValAlaProIleuLeuGIu 20
Db 91 ATGGCGGAGGAAAGACCGTGGCGGTGGTGGAGTGGCGGCTGAACACCAAGAA 150
QY 21 Gln-----GlyAspGlnAlaAsnLeuGlnIleProIleValGlaGlyAsnAnI 38
Db 151 GAATCATTGGAGAAACCTGCCAAGTTTACTGAAAACGACATATATGCTATTTCAA 210
QY 39 ValAspGIuThrLysSerPheAsnPhesApArgValPheAsnSerHisGluSerThr 58
Db 211 GTTGATGGAAGATTAATCCTTCATTTGATGCTGTCCTTTCATGTATGTAAGTACC 270
QY 59 GlnIleIyrgInGIuIleAlaValProIleIleIerSerAlaLeuGlnGIyIyAsnGI 78
Db 271 AATGGTATGGAAGAAATGACACACCAATCATCTGATTCGTCACAGGCTACATGT 330
QY 79 ThrIlePheAlaIyrgInGIuThrSerSerGIyIyThrIyThrMetGlyThrPro 98
Db 331 ACTATATTTGGCTATGACAGACAGTGGCTTCAGAAAACATATACATGATGAGGTTC 390
QY 99 AsnSerLeuGIyIleIleProGlnAlaIleGlnIleValPheLysIleIleGlnGIu 118
Db 391 GATCTTTTGGGAGTATATCCACAGGCAATTCATACATTTCCAAAATAATTAAGACT 450
QY 119 ProAsnArgGluPheLeuLeuArgValSerIyMetGluIleIyAsnGIuThrValLys 138
Db 451 CCTGATAGGAATTTCTCTTACGTGTATCTTACATGGAATAATACATGAACCATACA 510
QY 139 AspLeuLeuCysAspAspArgIyrgLysIySProIeuGlnIleArgGluAspPheAsn 158
Db 511 GATTTTACTGTGGCAGCTCAAAAATAATGAACCTTTAATATTATGAGAAGATGTC 570
QY 159 AsnValIyValAlaAspLeuThrIleGluIleuValMetValProGlnHisValIleGln 178
Db 571 AATGTGTATGTCGCATCTCCACAAAGCTGTATATACATCAGAAATGGCTTTAAA 630
QY 179 TrpIleLysGIuGIuLysAsnArgHisIyrgGIuThrLysMetAsnAspHisSer 198
Db 631 TGGATTACAAAGGAGAAAAGCAGCAGCATTTATGAGAAAACAAAATAATCAAGAAG 690
QY 199 SerArgSerHisThrIlePheArgMetIleValGluSerArgAspArgAsnProThr 218
Db 691 AGTCGTTCATACACATCTTTTNGATGATTTTGGAAAACAGAGAAGAGGCTGAAC 750
QY 219 AsnSerGIuAsnCysAspGIyAlaValMetValSerHisIleuAsnLeuValAspLeu 238
Db 751 -----AATGTGAAGATCTGTATGAAGTATCCATTTGAATTTGGTATCTTGA 801
QY 239 GlySerGIuArgAlaSerGIuThrIyrgIleGluGIuValATyLeuLysGIuGIyAsn 258
Db 802 GCGAGTGAAGAAGACTCTCAAAACAGCGCTGCAGGTGTCCGGCTCAAGAAAGCTGTAT 866

```

```

Qy 239 ILeaAsnArgSerLeuPheIleLeuGlnValIleLysLysLeuSerAspGlnAla 278
Db 862 ATAAATCAAGCCTTATTATTGGGACAAAGTACGAAGAACTTAGTGACAAAGTT 921
Qy 279 GlnGlyPheIleAsnTyrArgAspSerLysLeuThrArgIleLeuGlnAsnSerLeuGly 298
Db 922 GGTGGTTTCATAAAATATATCGAGATAGCAAGTACACGAAATTCCTGAGATTCCTGGGA 981
Qy 299 GlnAsnAlaLysThrValIleLeuCysThrIleThrProValSerPheAspGlnThrLeu 318
Db 982 GGAATTCACAAAGACACGATATTATCTGCACAAATTACTCCAGTATCTTTTGATCAAACTCTT 1041
Qy 319 SerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGln 338
Db 1042 ACTGCTTCCAGTTTGGCCAGTACTGCTAAATATATGAGAAATATCCCTTATGTTAATGAG 1101
Qy 339 ValLeuAspAspGlnAlaLeuLeuLysArgTyrArgGlyGlnIleLeuAspLeuLysLys 358
Db 1102 GTATCCAACTGATGAGAGCTCTCTGAAAAGTATAGAAAAGAAATATGATCTTAAAAAA 1161
Qy 359 GlnLeuGlnAsnLeuGlnSerSerSerGlnThrLysAlaGlnAlaMetAlaLysGlnGln 378
Db 1162 CAATTAGAG-----GAGGTTCTTTAGAGACGCGGCTCAGGCAATGGAAGAAACCA 1215
Qy 379 HisThrGlnLeuLeuAlaGlnIleLysGlnLeuHisLysGlnArgGlnAspArgIleTrp 398
Db 1216 TTGGCCCAACTTTTGGAGAAAAGATTGGCTCAGAAAAGTACCAATAGAGAAAATTGAA 1275
Qy 399 HisLeuThrAsnIleValAlaLysSerGlnGlnUser---GlnGlnAspGlnArgVal 417
Db 1276 AACTTAACACGAGTCTGTGCTCTCTCTCCCTCAGCTGCAACAGAAATTAAGGCT 1335
Qy 418 LysArgLysArgArgValThrTyrAlaProGlnLysIleGlnAsnSerLeuHisLysSer 437
Db 1336 AAAAGAAAACGACAGATTACTTGCTGCTTGGCAAAATTAACAAATGACAGAACTCAAC 1395
Qy 438 GlnValSerAspPheAspMetLeuSerArgLeuProGlnAsnPheSerLysLysAlaLys 457
Db 1396 TATGCGATCAATTAAAT-----ATACCAACAATATATACAAACAACAAACACAT 1443
Qy 458 PheSerAspMetProSerPheProGlnIleAspAspSerValCysThrGlnPheSerAsp 477
Db 1444 AAGCTTCTATAAATTATTAACGAGAAATGATGAATCTGCTGTGACAGCTGATGTT 1503
Qy 478 PheAspAspAlaLeuSerMetMetAspSerAsnGlnIleAspAlaGlnThrProAsnLeuAla 497
Db 1504 TTCAGTACACTCTGTTGATACATTAAGT-----GAGATGAAATGGAATCCAGCA 1551
Qy 498 SerLysValIleThrHisArgGlnLysThrSerLeuHisGlnSerMetIleAspPheGln 517
Db 1552 ACAAGCTACTAAATCAGAG----- 1572
Qy 518 IleSerAspSerValGlnPheHisAspSerSerLysGlnAsnGlnLeuGlnIleTyrLeuPro 537
Db 1572 ----- 1572
Qy 538 LysAspSerGlnAspMetAlaGlnCysArgLysAlaSerPheGlnLysGlnIleThrSer 557
Db 1572 ----- 1572
Qy 558 LeuGlnGlnGlnLeuGlnSerLysGlnGlnLysLysGlnLeuValGlnSerPheGln 577
Db 1572 ----- 1572
Qy 578 LeuLysIleAlaGlnLeuGlnGlnGlnLeuSerValLysAlaLysAsnLeuGlnMetVal 597
Db 1573 ----- 1587
Qy 598 ThrAsnSerArgGlnHisSerIleAsnAlaGlnValGlnThrAspValGlnLysGlnVal 617
Db 1588 TTGAACCTA----- 1596
Qy 618 ValArgLysGlnMetSerValLeuGlnLysAspSerGlnTyrAsnAlaSerAsnSerAspLeu 637

```

```

Db 1596 ----- 1596
Qy 638 GlnAspSerSerValAspGlnLysArgLeuSerSerSerHisAspGlnLysIleGlnHis 657
Db 1596 ----- 1596
Qy 658 ArgLysMetLeuGlnGlnLysIleValAspLeuGlnGlnPheIleGlnAsnLeuAsnLys 677
Db 1596 ----- 1596
Qy 678 LysSerGlnAsnAspLysGlnLysSerSerGlnAsnAspMetGlnSerIleGlnLeu 697
Db 1596 ----- 1596
Qy 698 CysGlnAlaIleMetAlaGlnLysAlaAsnAlaLeuGlnGlnLeuAlaLeuMetArgAsp 717
Db 1597 ----- 1605
Qy 718 AsnPheAspAsnIleIleLeuGlnAsnGlnThrLeuLysArgGlnIleAlaAspLeuGln 737
Db 1606 GACTATGATATCTGGTATTACATATGAAACAACATACGACGAAAGAAAGAAATGGA 1665
Qy 738 ArgSerLeuLysGlnAsnGlnGlnThrAsnGlnPheGlnIleLeuGlnLysGlnThrGln 757
Db 1666 TTGAATTTAAAGAAAAGAAATGATTTGATGAAATTTGAGAGCTGTAAGAAAGAAACTAA 1725
Qy 758 LysGlnHisGlnAlaGlnLeuIleHisGlnIleGlySerLeuLysLysLeuValGlnAsn 777
Db 1726 AAAGATCAAGAGATCAACAAATTAATCATGAAATTTGCAAAATTAATTAATTAAGAT 1785
Qy 778 AlaGlnMetTyrAsnGlnAsnLeuGlnGlnLysAspLeuGlnThrLysThrLysLeuLys 797
Db 1786 CGAGAAATATATATATCAAGATCTTGAGATGAACTACAGTCAAAAGTACAGCTGCTTAA 1845
Qy 798 GlnGlnGlnIleGlnLeuAlaGlnLeuArgLysArgAlaAspAsnLeuGlnLysLysVal 817
Db 1846 GAAAAGAAAGACAGATTAAGAACTACAGAAATACATAGACTCAAAAGCTAAGAAAT 1905
Qy 818 ArgAsnPheAspLeuSerValSerMetGlnYAspSerGlnLysLeuGlnLysIlePhe 837
Db 1906 ATAAAGATGACTTGCATCACTCATTTG-----GAAAGCATTTGAACACCAAAA 1953
Qy 838 GlnLeuLysGlnSerLeuSerAspAlaGlnAlaValThrArgAspAlaGlnLysGlnLys 857
Db 1954 CAATGAAGCAGACGCTGTTGATGCTGAAGTACAGCCCTGATGACCAAGAGAAATCA 2013
Qy 858 SerPheLeuArgSerGlnAsnLeuGlnLysLysGlnLysMetGlnAspThrSerAsnTrp 877
Db 2014 GCCTTTCTTAGAAGTAAATCTGAGTTGAAAGGAGAAATGAAGAACTTCAACTACA 2073
Qy 878 TyrAsnGlnLysGlnLysAlaAlaSerLeuPheGlnLysGlnLeuGlnThrLysLysSer 897
Db 2074 TACAAGCAAAATGGAATGATATTCAGTTATATCAAGCCAAATTTGGAGCAAAA----- 2127
Qy 898 AsnTyrLysMetGlnAlaAspLeuGlnLysGlnLeuGlnSerAlaPheAsnGlnIle 917
Db 2128 -----AAGAAATTCAGATTGCTGGAGAAAGAAATTAACATTCGCTTTAATGAGATA 2181
Qy 918 AsnTyrLeuAsnGlnLysLeuAlaGlnLysValProArgAspLeuLeuSerArgValGln 937
Db 2182 ACAAAACCTCACCTCCCTATATAGTGCAGAAAGTTCCAAAAGATTTCTCTGTAATTTGGA 2241
Qy 938 LeuGlnLysLysValSerGlnPheSerLysGlnLeuGlnLysAlaLeuGlnLysAsn 957
Db 2242 TTGGAAAGAAAGATTACGATCTTCAGAAAGAACTTAATTAAGAAATGGAAGAAATGAA 2301
Qy 958 AlaLeuGlnAsnGlnValThrCysLeuSerGlnTyrLysPheLeuProAsnGlnValGln 977
Db 2302 GCTTTGGCGGAGAAAGATCATTTGCTTTCAGAAATTAATTTACCTTCTGAAATAGAA 2361
Qy 978 CysLeuLysAsnGlnIleSerLysAlaSerGlnGlnIleMetLeuLeuGlnGlnLys 997

```


Db 2362 AGGCTGAGAAAGATGACAGCAAAATCTGAAGAGCTCCATATTAATACATGAGAAAA 2421
QY 998 GluHisSerAlaSerIleIleSerLysGluGluIleMetGlnGlnIleSerGluGln 1017
Db 2422 GATTAATTTGTTTCTGTAAGTAGTTCAATAGAGAGTAGAGTAGAGTTACTTAAAGAA 2481
QY 1018 IleGlnIleuThrAspGluValIleHisIleThrGlnSerLysValGlnIleThrGln 1037
Db 2482 ATTTGGGAAAACAAAGATGACCTACACATACACAGTCGATTAATTAAGACATGATCAA 2541
QY 1038 GlnIleuGlnIleuMetLysMetHisAspAspLeuPheGlnLysTyr 1053
Db 2542 GAATTCGAAATTTCAAAACCCCTCATATGACTTTGACCAAAAGTAAATGATGTCCTT 2601
QY 1054 -----IleArgAsnLysSerGluAlaGlnAspLeuArgGluMetGluAsnLeu 1070
Db 2602 GAGGAGATGAGAGAAATGAAATCAGGAAATAGTTATCTCTTAAGAAAGCCCAAAATTT 2661
QY 1071 LysGluThrMetGlnSerValGluValLysIleAlaAspThrLysHisGluLeuGlu 1090
Db 2662 GATTCGAGTTGGGTGCTTGAAGACGAGCTTTCTTACAAGACCCCAAGAACTTGAGAG 2721
QY 1091 ThrIleArgAspLysGluGlnIleuLeuHisGlnLysLysTyrPhePheGlnAlaMetGln 1110
Db 2722 AAAACAGTGAGTTCAAGAAAGACTAAATGAGATGAGACAGCTGAAGGAAACATTTGAA 2781
QY 1111 ThrIlePheProIleThrProLeuSerAspSerLeuProProSerLysLeuValGluGly 1130
Db 2782 AAT-----AGAGATTCTCCCTCAAACTGACAAAGGAGAAACACGATTTACGAG 2835
QY 1131 AsnSerGlnAspProIleGluIleAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGlu 1150
Db 2836 AAACGTGACGCAACATTTAGAA-----GAAGTAAAACTTAACTCAAGAAAGAAAGAT 2889
QY 1151 ArgAsnAsnIleMetValCysLeuGlnIleThrGluArgAsnSerLysGluGlnVal 1169
Db 2890 CTAAACAACTCCAAAGAAAGCTGCAATTTGAGAGGACCACTCAAAAGATGATTTAC 2949
QY 1170 -----IleAspLeuAsnThrGlnLeuGln-----SerLeuGlnAlaGln 1182
Db 2950 GATACCTGTAACATGAAATATGATATCATCAAGACATTAACGAAATGCTTGAATCTCTG 3009
QY 1183 SerIleGluLysSerAspLeuGlnLysProLysGlnAspLeuGlnGluGluValLys 1202
Db 3010 AAACAACTCAAGAAACAAATTAATACACTAAATCGAAATTTCTGAG--GAAGTTTCC 3066
QY 1203 LeuLeuLeuGlnMetGlnIleLeuLysGlnIleHisLeuThrAsp-----SerGln 1218
Db 3067 AGCAATTTGCAATGAGAGGAAATATCAGAGAGAACTAAAGATGATTTCAAGCAAAAGATG 3126
QY 1219 LeuSerIleGluLysLeuGlnIleuGlnAsnLeuGlnValIleThrGluLysLeuGlnThrLeu 1238
Db 3127 GTTGGCATGATTAATAAAAAACAG-----GATTTGGAA--GCTAAATAATACCAAAACCTA 3177
QY 1239 GlnGluGlnMetLysAsnIleThrIle----- 1247
Db 3178 ACTGAGATGTTAAAGGATTAATGAGATTAATGAGCAAAAGAGATATTTCTTTAATA 3237
QY 1248 ---GluArgAsnGlnLeuGlnIleThrAsnPheGlnAspLeuLysAlaGlnHisAspSerLeu 1266
Db 3238 CAGGAGAAATGAACTCCAAACAAATGTTAGAGAGTGTATATACGCAAGAAAGAAACAATG 3297
QY 1267 LysGlnAspLeuSerGlnAsnIleGlnIleuSerIleGluThrGlnAspGluLeuArgAla 1286
Db 3298 AAGACTGACCTTAAGCAAAATATGTAATGACATTTGAAACAGAGAAATTAAGACTT 3357
QY 1287 AlaGlnGlnIleuLeuArgGlnIleuLysGlnIleuValAspSerPheArgGlnIleuLeu 1306
Db 3358 CTGGGGATGAACTTAAGCAAGACAAAGATGTTGCAAGAAAGAAAGAACCATGCCATA 3417
QY 1307 AspCysSerValGlyIleSerSerProAsnHisAspAlaValAlaAsnGlnIleLysVal 1326
Db 3418 AAG---AAGAAAGAGAGACTTTCTAGACACTGTGTACAGACTGGCAGAAATTTGAAGAAAA 3474

QY 1327 SerLeuGlnValAlaAsnSerLeuGlnSerGlnMet-----LeuArgGly 1341
Db 3475 CTAAAGGAAAGAAAGCCAGCACTCCAGAGAAAGAAACGAAACAACTTTAATGATCAAGAA 3534
QY 1342 GluArgAspGlnLeuGlnIleThrSerCysLysAlaLeuValSerGlnLeuGlnLeuArg 1361
Db 3535 GAGATGAGTGAAGATGCGAAAAAG-----ATTAAAGAAATAGAGAAATTTAAAG 3582
QY 1362 AlaHisValLysSer-----ValGlnGluGlnAsnLeuGlnIle 1374
Db 3583 AATGAATTAAGAGCAAAACAAATTTGACATTTGCAACATATGCAAGAAACAGAGGCTTAGTTG 3642
QY 1375 ThrLysLysLeuAsnGlnLeuGlnLysGluIleLeuGlnLysSerGlnIleSerGluVal 1394
Db 3643 GCTCAAGAACTTAATGAAATTAATGAGAGAGTGAATCTTAACCAAGAAAGAAAGAT 3702
QY 1395 LeuLysSerMetLeuGlnLysAsnLeuLysGlnAspAsnAspLysLeuLysGlnIleAlaGlu 1414
Db 3703 CTAAAGGAATTTACAGAAAGTCATTGAAACAGAGAGACACCTTAGAGAGATAT----- 3756
QY 1415 GluTyrSerSerLysGlnAsnGlnPheSerLeuGlnGluValPheSerGlnSerGlnLys 1434
Db 3757 -----ATAGAGAAATTTGAAGCTACAGGCTTACAA 3786
QY 1435 LeuValAspGluIleGluValLeuLysAlaGlnLeuLysAlaAlaGlnLysArgLeuGlu 1454
Db 3787 ACCAAAGAGAAAGAACTTAAATTTGCTCATATTCACCTTAAAGAACACCAAACTATTTGAT 3846
QY 1455 IleLysAspArgAspTyrPheGlu--LeuValGlnIleThrAlaAsnThrAsnLeuValGlu 1473
Db 3847 GAACATAAGAGAGAGCGTATCTGAAAGACAGCTCAAAATTAATATCTCAGAGCTTAGAA 3906
QY 1474 GlyLysLeuGlnIleThrProLeuGlnAlaAsp-----HisGlnLysLys 1487
Db 3907 ---AAATCCATGACCAATTAAGAAAGAGATCCAGCTGCTTCAAGCAACAAAGATTA 3963
QY 1488 -----SerIleAspArgArgSerGlnGlnMet--GluLysValLeu 1501
Db 3964 CTGCTTAATGTGAAGAAAGTCACTGAGACTCAGAGAAACAAATGATGACGAGATTAATA 4023
QY 1502 GlyLysLys---LeuGlnArgAsnGlnTyrLeuLeuGlnLysArgGlnGlnLysLeu 1520
Db 4024 ACAGAACAGTCCAAACCAAGAGACTCAGAACACTGGCAAGATAGAAATGAGAAAGCTC 4083
QY 1521 GluLeuSerAsnLysLeuGlnIleLeuGlnLysGlnMetGluThrSerValLeuLys 1540
Db 4084 AGGTGAAATGAAAAATTTCAAGAAAGTCAAGAGAGATTAATCTTAACCAAGAAAGAA 4143
QY 1541 AspAspLeuGlnGlnLysLeuGlnIleuSerLeuSerGlnAsnIleIleLeuLysGlnAsn 1560
Db 4144 GACAACTTAAACGATTAAGAGAGCCCTTGAAAGTTAAACATGACACAGCTGAAGAAAGAT 4203
QY 1561 IleAspThrThrLeuLeuLysHisHisSerAspThrGlnAlaGlnIleuGlnLysThrGlnGln 1580
Db 4204 ATTGAGAAACTTTG-----GCTAAATTCAGAGCTCAAGAC 4242
QY 1581 GluLeuGlnIleuAlaLysAsnLeuAlaIleAlaLysSerAspAsnCysProIleThrGln 1600
Db 4243 AAACAAAGAACAGTCCCTTAATATGAAAGAAAGAAAGACATAAACTCAACAAAGCTGAGT 4302
QY 1601 GluLysGln-----ThrSerAlaAspCysValHisProLeuGlnGlnLysIleLeu 1617
Db 4303 GAGATGAGCAATTAACCAACCAAGATTCAGACACTACTAAGGATAGAAATAGAAATGCTC 4362
QY 1618 LeuLeuThrGlnGlnIleuHisGlnLysThrAsnGlnGlnLysLeuLeuHisGlnLys 1637
Db 4363 GGATTTGCCAAAGAGCTTCAAGAAAGTCAAGATGATGAATGAAATCTGATGAGAGAGAA 4422
QY 1638 AsnGlnLeuGlnIleGlnAlaGlnValGlnLysCysGlnValGlnHisLeuMetLysSer 1657
Db 4423 GATGACCTACAGAGGCTTCAAGAAAGATTTCTCAATCTGAAAGTACAGCTC----- 4473

Db 6480 ----- 6480
 QY 2320 GlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPheLysVal 2339
 Db 6481 -----ATCATGAAGAAACGCAAGTAT 6501
 QY 2340 Val-----TyrArgThrAlaAlaValLysGlnAspHisSerLeuLysAspTyrGlu 2357
 Db 6502 GTCTTAAGCATGTTCAAAATAAAGAAACATCATGATCATTAATTAATTGAA 6561
 QY 2358 LysAspLeuAlaIleGlnGlnLysArgHisAspGlnLeuArgLeuGlnLeuGlnCysLeu 2377
 Db 6562 ATGGATTTTGTGATGAGTGGAAAGCAAAAGCAATGCTAATTAATAATCAACCTT 6621
 QY 2378 GlnGlnHisGlyArgLysTyrSerAspSerAlaSerGlnLeuLysPheCysGluIle 2397
 Db 6622 CAACAAGAT-----TGTGATGTCACATCCAGCAATTAAGGATCTCAATTG 6669
 QY 2398 GluPheLeuAsnGlnLeuLeuPheLysLysAlaAsnIleIleGlnSerValGlnAspAsp 2417
 Db 6670 AACCAAGATATGATCTA-----CAATGTGAGAAATTCCTCAAGAT 6711
 QY 2418 PheSerGluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlnLeuGlu 2437
 Db 6712 TTCTCGAAGAAAGTGAG-----TTCCCTAGCATTAAGACTGAATTCACAAAGTACTAAGT 6765
 QY 2438 HisLysLysGlyPheMetGlnTyrPheGlnGluGluPheGlyAspLeuHisValAspAlaLys 2457
 Db 6766 AATAGGAAGAAATGACACAGCTTTTGGAGAGTGGTTAATCTGCTTTGATATATGAA 6825
 QY 2458 LysLeuSerGlnGlnMetGlnGlnGlnLysAsnArgArgIleAlaSerThrIleGlnLeuLeu 2477
 Db 6826 AACCTTAATAATGGCATCCGAAAGAAATGATAGGATTTGTCAAGTAACTCTT 6885
 QY 2478 ThrLysArgLeuLysAlaValGlnSerLysIle-----GlnArgGluIleThr 2494
 Db 6886 AATAACAGATTAATTCCTCAATTAATGAATGAATCAACAGAGTTTGAGCAAGAAAGTGCTACC 6945
 QY 2495 ValTyrLeuAsnGlnPheGlnAlaLysLeuGlnGlnLysGlnGlnAsnLysGlnLeu 2514
 Db 6946 ATA-----TCCAAAGAGTGGCAAGACGACTGAATTCACCTGAAGCAAAATGAAAACTA 7002
 QY 2515 MetArgArgMetGlnHisHisGlyProSerAlaSerValMetGlnGlnGlnAsnAlaArg 2534
 Db 7003 TTTAAAACTACCAACATTTGAAGACTCTCTGTGGCATCTGTGCCAGGTTAAT----- 7056
 QY 2535 LeuLeuGlyLysLeuLysThrValGlnAspGlu-----SerLysLys 2548
 Db 7057 -----CCTACCAACAAGACAAATGAATCCTCATGTTACATCAAGAGCT 7101
 QY 2549 LeuGln-----SerArgIleLysMetLeuGlnAsnGlnLeuAsnLeuValLysAsp 2565
 Db 7102 ACACAGATTAAACAGAGAAATTCGAGACTCGAAATTCACCTGAAAGCTGAAGAA 7161
 QY 2566 AspAlaMetHisLysGlnGlnValAlaIleLeuGlnAspLysLeu-----LeuSerArg 2584
 Db 7162 AGTGCATGATGATAGAAAGCAAGATTAATAAGATGCAGAAAGACTGTGAGGACTAAT 7221
 QY 2585 AsnAlaGlnLysGlnLysLeuAsnAlaMetGlnValLysLeuThrLysLysGlnAspAsnLeu 2604
 Db 7222 GACATAATAGCAAACTT-----CAAGCCAAAGTTTATGATCAATTAATGCTTT 7272
 QY 2605 GlnAlaAlaMetLysGlnIleGlnLeuAsnLeuGlnLysMetValAlaLysGlyAlaValPro 2624
 Db 7273 GAAAAAACAAGACAGACATTAAGTACTTCAGAGACAAGTCTTTAGAGAGTAAAGCCA 7332
 QY 2625 TyrLysGlnGlnIleAspAsnLeuLysThrLysValValLysIleGlnMetGlnLysIle 2644
 Db 7333 TATTAAGAAGAAATTAAGATCTCAAAATGAACTGTGAAAAATAGACCTAGAGAAAGT 7392
 QY 2645 LysTyrSerLysAlaThrAspGlnGlnIleAlaLysLysSerCysLeuGlnAspLys 2664
 Db 7393 AAAAATGCCAAAGAAATTTGAAAAAGAAATCAGTCTACAAAGCCACTGTAGATATATCA 7452

QY 2665 GlnGlnLysLeuArgArgLeuLysGlnGlnLeuArgArgIleAlaAspAsnAspThr 2684
 Db 7453 AAGGAAGTTATTAAGCGCATTTAGAGAAATCTCGAAGAAAGTCAACAGGCCCAAGATACC 7512
 QY 2685 ThrValCysValProLysAspTyrGlnLysAlaSerThrPheProValThrCysGlyGly 2704
 Db 7513 TCAAGTATATACAGAACATCTGATCCACGCTTCAATTAACCTTAATCTGTGGAGT 7572
 QY 2705 GlySerGlyIleValGlnSerThrAlaMetLeuValLeuGlnSerGlnLysAlaLeu 2724
 Db 7573 GGCAGCGGCATTTGACAAACACAAAGCTTATTTGAAAAGTGAACTATTAAGGCTA 7632
 QY 2725 GlnArgGlnLeuSerHisTyrLysLysTyrHisHisLeuSerArgThrMetSerSer 2744
 Db 7633 GAAAAAAGAAATTTCTAAGTTAAAGCAGCAAAATGACAGCTA----- 7674
 QY 2745 SerGlnAspArgLysLysThrLysAlaLysSerAspAlaHisSerSerHisThrGlySer 2764
 Db 7675 -----ATMAAACAAAGAAATGAAATTTTAAGCAATTAATCAGCATCTT 7716
 QY 2765 SerHisArgLysSerProHisLysThrGluThrTyrArg-----HisGlyProVal 2781
 Db 7717 TCCAAATGAGTCAAAACTTGGAAAGAAAGAACCTTAAGAGAGAGCTCAAAACAAAGTA 7776
 QY 2782 ThrProGlnArgSerGlnMetProSerLeuHisLeuGlySerProLysLysSerGluSer 2801
 Db 7777 ACTGTGAG-----AATCTCCAAAG----- 7797
 QY 2802 SerThrLysArgValAlaSerProAsnArgSerGluIleTyrSerGlnLeuValMetSer 2821
 Db 7798 -----TCT 7800
 QY 2822 ProGlyLysThrGlnMetHisLysHisIleLeuSerProSerLysValGlyLeuHisLys 2841
 Db 7801 CCTAAAGTGACTGAGACA-----GCTTCTAA 7827
 QY 2842 LysArgAlaLeuSerProAsnArgSerGlnMetProThrGlnHisValIleSerProGly 2861
 Db 7828 AAGAAACAATTTACACC----- 7845
 QY 2862 LysThrGlyLeuHisLysAsnLeuThrGlnSerThrLeuPheAspAsnLeuSerSerPro 2881
 Db 7846 -----TCTCAA 7851
 QY 2882 CysLysGlnGlnLysValGlnGlnLeu-----AsnSerProLysGlyLysLeuPhe 2899
 Db 7852 TGCAGGAAGAGAAATTTACAAAGATCCTGTGCCAAGAAATCACCAAAATCTGTTTTT 7911
 QY 2900 AspValLysSerLysSerMetProTyr-----CysProSerGlnPhePheAspAsnSerLys 2918
 Db 7912 GATAGCCGATCAAAAGTCTTATACATCACCTCATCCAGAGTTCGATTTTGTATCTCAAG 7971
 QY 2919 LeuGlyAspPheSerGlnLeuAsnThrAlaGlnSerAsnAspLysSerGlnAlaGlnAsn 2938
 Db 7972 TTAGGCGCTTTGCCAGAGGTGCAAAATGCAAGAGAGAGTGTGATCTCAGCCAGGT 8031
 QY 2939 TrpTyrGlnAlaLysLysGlnThrAlaProGlnCysLysThr 2953
 Db 8032 CCTTGGCAGCCCTCTCAGGCAAGATGTCCGTGATGCAAACT 8076
 RESULT 3
 US-09-960-253-145
 ; Sequence 145, Application US/09960253
 ; Patent No. US20020123619A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Lodes, Michael J.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.556
 ; CURRENT APPLICATION NUMBER: US/09/960,253

```

: CURRENT FILING DATE: 2001-09-20
: NUMBER OF SEQ ID NOS: 187
: SOFTWARE: fastseq for Windows Version 4.0
: SEQ ID NO 145
: LENGTH: 10300
: type: DNA
: ORGANISM: Homo sapiens
: OS-09-960-253-145.

```

Pred. NO.:	1.01e-64	Length:	103000
Score:	1153.00	Matches:	675
Percent Similarity:	39.53%	Conservative:	651
Best local Similarity:	20.13%	Mismatches:	1170
Query Match:	7.81%	Indels:	858
DB:	10	Gaps:	138

US-09-150-867-1 (1-2954) x US-09-960-253-145 (1-10300)

OY	110	GlValPheLysIleIleGlnGluIleProAsnArgGluPheLeuLeuArgValSerTyr	129
Db	559	GAATTTGAAAAAGTAAACATAAGCTCCAGACAGAAAGCAAACTATATC-----	606
OY	130	MetGluIleTyrAsnGlnIuThrValIysAspLeuLeuCysAspAspArgTyrGlySerPro	149
Db	607	-----AGCACTTTGCAAGCCCAAGCTTACTACGACAGCAGCAGCAACAACT	651
OY	150	LeuGluIleArgGluAspPheAsnArgAsnValTyrValAlaAspLeuThrGluLeu	169
Db	652	GCACAGAGTCTCTACAGAGAG-----GAAAGAAATT	681
OY	170	ValMetAlaProGluHisValIleGlnIleThrIleLysGlyGluIuLysAsnArgHisTyr	189
Db	682	GTAAATGATGAGACACAGCTC-----CAGAGAAAGCAAAATTCATT	723
OY	190	GlyGluIuThrLysMetAsnAspHisSerSerArgSerHisIleThrIlePheArgMetIleVal	209
Db	724	AGCACTTACAAACCCAGCTCAGCCAGACACAGACAGCAAGCAAGCTGCACAGCAGGTGTC	783
OY	210	GluSerArgAspArgAsnAspProThrAsnSerGluAsnCysAspGlyAlaValMetVal	229
Db	784	CGAGACAGAAAGATCCCGCTTTGAAACAAGTTCGCTTCATGAAGATGAGCTT-----	837
OY	230	SerHisLeuAsnLeuValAspLeuAlaGlySerGluIuArgAlaSerGlnIleThrIleLysGlu	249
Db	838	-----CTTCAGTTAGTAAACCCAGCAGATGTGGAAACAGAGATGCA-----	879
OY	250	GlyValArgLeuLysGluGlyCysAsnIleAsnArgSerLeuPheIleLeuGlyGlnVal	269
Db	880	-----CAGAAATTGAGGTGCTGCAGAAAGAG	906
OY	270	IleLysLysLeuSerAspGlyGlnAlaGlyGlyPheIleAsnTyrArgAspSerLysLeu	289
Db	907	CTTGAGAGAACAGCAAGAACTCTGTGGTGGC-----CGTCTCAGGTG	948
OY	290	ThrArgIleLeuGlnAsnSerLeuGlyGlyAsnAlaLysThrValIleIleCysThrIle	309
Db	949	GTTGACTTCTGCACAGAGAGCTGACTGCTGTGACAGACAGA-----	990
OY	310	ThrProAlaSerPheAspGluIuThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHis	329
Db	991	-----AACCGATTCCTCTCCAGCACTTACAGCAGACAGATGAGCACTCAGCAT	1035
OY	330	-----ValArgAsnThrProHisValAsnGluValIleAspAspGluAlaLeuLys	347
Db	1036	AATACTTTGAGGAACACTGTGGAA-----ACAGAAAGAGAGAGATCCCAAGTTTACTGGAA	1092
OY	348	ArgTyrArgLysGluIleAsnAspLeuLysLysGlnLeuGluAsnLeuGluIuSerSerSer	367
Db	1093	AAGATGAACTTGAGTGGCAGAGAAATAATATCTCTCCATATCTGCAAGAAAGAAATG	1155
OY	368	-----GluThrLysAlaGlnAlaMetAlaLysGluGluIuHisThr	380

Dd	1153	CATGATCTTTTAAAGAAAGSTTTGAGCAAGGAGGCGCAAGGCCAGGCGGAGGCTGAACTAGAG-----	1208
Oy	391	GlnLeuLeuAlaGluIleLysGlnLeuHisLys--GluArgGlnAspArgIleThrHis	399
Dd	1207	TCTGGTATAGTCTGTTTGGAGCAAGAGCAAGAAAGCAAGAAATGAAAGAAAGAGAGCTCTGAT	1266
Oy	400	LeuThrAsnIleValValAlaSerSerGln	409
Dd	1267	ATTGTGAGCTCTCAAAAGAGCTGCGACAGAGAGCTGAGTCTGCTGTGATCTGTAAAGAT	1336
Oy	410	-----GlnSerGlnAspArgValValuArgVal	423
Dd	1327	CAAAATTCAGAGCTCTCCAGATAGATAGATAGAGAGAGCTTCAAGTCAAGCCAGGAGCAT	1386
Oy	424	ThrThrAlaProGluLysIleGlnAsnSerLeuHisAlaSerGluValSerAspPheAsp	443
Dd	1387	CAGAGAGCTGGAAGATGAGCTGCGACAGCAAAA-----TCCAAAGAAATGAGCCAAATT---	1437
Oy	444	MetLeuSerAlaGluProGluAsnPheSerLysValValLysPheSerAspMetProSer	463
Dd	1438	--STAATTAACAGTCCGCTCTTCAACCAACATGAAGAAACAGCACTGAGACTCTCT	1489
Oy	464	PheProGluIle-----AspAspSerValLysThr	473
Dd	1486	TTCCCAAGATGTTTAAATGAGAGGCGACAGGCGAGCTGAGAGAGATTTCTCTTCTTGG	1553
Oy	474	-----GlnPheSerAspPheAspAspAla-----LeuSerMetMetAsp	486
Dd	1546	CAGAAAGAGAGTGGTAAAGATAGAAATGAAGAAAGAGAGCTGTGCTGTATGATAG	1605
Oy	487	SerAsnGlyIleAspAlaGlu---ThrAsnLeuAlaSerLysValThr-----	501
Dd	1606	CTGGAGGAGCTGAAGAGCTGAGAAATGAAGAACTCTCTTCAGATTACTCTCTAGAGCT	1665
Oy	502	-----HisArgGlnLysSerLeuIleGlnSerMetLeu	514
Dd	1666	CAGATTAGAGCTGGGAGGAGCGACAGACAGAGAGTGAAGTGAAGATC-----AGCATTTGTAT	1719
Oy	515	Phe-----GlyGlnIleSerAspSerValGln	523
Dd	1720	ATTGGCAACAGAGAGAGCTGTCTGCTGAGAAAGTGGCAAA-----GATCTCTAGAA	1773
Oy	524	PheHisAspSerSerLysGlnAsnGlnLeuGlnIleLysProLysAspSerGluAspMet	543
Dd	1774	AACACATTTTCTCGAAACACATTAAGATTTATCAGTTTAA-----TTG	1815
Oy	544	AlaGlnCysArgLysAlaSerPheGlnLysGlnIleLeuSerLeuGlnGlnIleGln	563
Dd	1816	TTGGAAATGAAGAGAGCT-----CAAGAGAGAAATGCACTTTCTTAAATTAAGAGCTCAG	1868
Oy	564	SerLys-----GlnGlnGlnLysLysGlnLeuValGlnSerPheGlnLeuLysIleAla	581
Dd	1870	GGAAGAAAGGCGCTGAGAAAGCAGATCATGAGAGCTCTTGACAGAAAGAAATGAAGACAGATG	1929
Oy	582	GlnLeuGlnGlnLysLeuSerValLysAlaLys-----Asn	593
Dd	1930	GAGGCTGAGGGAATAGCTCCCAATTAATGAAGATTTTCTTGAAGATCACAGGCGCAAGAT	1989
Oy	594	LeuGlnMetValThrAsnSerArgGlnHisSerIleAsnAla-----GluValGlnThr	611
Dd	1990	TTTCCCTTAATGCGCAAAAT---GAAGAGAGCAGCTCTTCCAGAGCTGAAAGAAACAGAGCGC	2046
Oy	612	AspValGlnLysGlnValValAlaArgLysGlnMetSerValLeuGluLysAspSerGluThrAsn	631
Dd	2047	AGCAGTGAACATCAAAATAGACATCTGAGAGAAATCTTTAAATAGATCTGAGATA---	2109
Oy	632	AlaSerAsnSerAspLeuGlnAspSerSerValAspGlyLysArg-----	646
Dd	2104	-----GAATTGAATCAACCAAGAGAGATGATGAATATCCCTTCTGCTGTA	2151
Oy	647	-----LeuSerSerSerHisAspGlnLysIleGlu-----	656
Dd	2152	CCAGATATTTGGCTAGTCTGCTCAGAGATGAGATGGAAGAGCTTAAAGTCAATTTTGGAG	2211

QY 657 -----HisArgLysMetLeuGlnGlnLys--- 664
 Db 2212 CTCGAGCTAAACTTTCATTAAGCACAGAAATCTATGAGAAAATTTAGATGAGAAAGCT 2271
 QY 665 -----LLeuValAspLeuGlnGlnLysValAspLeuSerGlnLysAsp 682
 Db 2272 AAGGAATTTAGCAACCTTAACCCAGTTGATTTGAGAGATTAAAGAAAATGCTCAGCAACAAC 2331
 QY 683 LysGlnLys-----SerSerGlnAspPheMetGlnSerLleGlnLeuLys 698
 Db 2332 AGCAGTGCATTCAGCTCTTTTGTCTGAGAGAAAGACAGCTTCTCTC---CAGGTGAAG 2388
 QY 699 GluAlaLleMetValGlnLysAlaAsnAla---LeuGlnGlnLeuAlaLeu----- 714
 Db 2389 GAACTTACATGTGTAACAGAAATTGAGGCTCAGTAAAGCAACTGGAATGACCTTGCA 2448
 QY 715 -----MetArgAspAsnPheAspAsnLleLe 723
 Db 2449 GAAAGCAAAAGCAAGAAAGAACTGTATTGAAAGCCAAACTGCCCATGACAACTCTCTC 2508
 QY 724 LeuGlnAsn-----GlnThr 728
 Db 2509 ACTGAACAGATCCATATGCTCAGCATAGACCCAAATCTAAAGATGTGAAAATTGAACTT 2568
 QY 729 LeuLysArgGlnLleAlaAspLeuGln-----Arg 738
 Db 2569 TTACAGAAATGAACTGATGATGATGCTCAGCTTCTGAGCAGAGTACCCGTATAGA 2628
 QY 739 SerLeuLysGlnLysAsnGlnGlnThrAsnGlnPheGlnLleGlnLysGlnLysGlnLys 758
 Db 2629 AGCTCGCAAAACCCAGCTCCAAATAAAGAAAGTGAAGTGTGAGGGCGCAACCTGTGA 2688
 QY 759 GluHisGlnAlaGlnLysLleHisGlnLleGlnSerLeu-----LysLysLeuValGln 776
 Db 2689 AGGATATCTCTCAAGTAAAGTGAAGAACTGTCCAGGCTCTTTCACAGAAAGCACTTGAA 2748
 QY 777 AsnAlaGlnMetLysArgAlaAspAsnLeuGlnGlnLys-----AspLeuGlnThrLysThrLys 794
 Db 2749 ATAACAAAATGATGATCAGCTTCTTACTAGAGAAAAGAGATGTGAAACCCCTCCAAACA 2808
 QY 795 LeuLeuLysGlnGlnGlnLysLleGlnLeuAlaGln----- 805
 Db 2809 ACCATCGAGAGAGATGATCAACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2868
 QY 805 ----- 805
 Db 2869 GTTCAGCTTAATGAAGAAAGTTTCTCTGGGGTTGAATTAAGACTTTAAGAACAG 2928
 QY 806 -----LeuArgLysArgAlaAspAsnLeuGlnLysValArgAsnPheAspLeuSer 823
 Db 2929 CTAAATTTATTTATCCAGAGCTGAGAGAAACAAAAGACAGAGTGAAGAAATATGAA 2988
 QY 824 ValSerMetLysAspSerGlnLysLeuGlnLysLlePheGlnLeuLysGlnSer--- 842
 Db 2989 GTTCTCTCTGCGC-----CTTAAACAAAATTAAT 3015
 QY 843 -----LeuSerAspAlaGlnAlaValThrArg---AspAlaGlnLysGlnLysSerPhe 859
 Db 3016 GATGAGATGAGCCGACAGAGCAAAATTAAGAAAGAACTTCAGCATGAAATTTGACCTT 3075
 QY 860 LeuArgSerGlnLysAsnLeuGlnLysGlnLysMet----- 871
 Db 3076 CTGAGAAAGAAATATGAGCAGAGAAAGAAAGAACTTCAGCAGCTCTTATTAAACAGAAAG 3135
 QY 872 -----GluAspThrSerAsnTrpLysAsnGlnLys 881
 Db 3136 GACCTTCTGCAAAAGCTCAGTATGAGAAAGAAATTAAGCAACTTGAAGATGATCT 3195
 QY 882 GluLysAlaLysSerLeuPheGlnLys-----GlnLeuGlnThrGlnLysSerAsn 898
 Db 3196 AAGAAAGAAATCCCACTCAGTGAAGTGAAGGGGAGAGAGTGAAGAAATTAAGAAAG 3255

QY 899 TyrLysLysMetGlnAlaAspLeuGlnLysGlnLysGlnLysSerAlaPheAsnGlnLysAsn 918
 Db 3256 AAGAAATCTCAGAA-----AAATGTGACTTCTAAGTGGCAAGAAATGAA 3303
 QY 919 ---TyrLeuAsnGlnLysLeuAlaGlnLys-----ValProArgAsp 931
 Db 3304 ATTTATTTAAACAGACATATCTGAGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3363
 QY 932 LeuLeuSerArgValGlnLeuGlnLysLysValSerGlnPheSerLysGlnLeuGlnLys 951
 Db 3364 TTGGAAGAAAGTGGCAGCTGAGAGCAATTCAGGCTGTGCAAAAGATGAATCAG 3423
 QY 952 AlaLeuGlnGlnLysAsnAlaLeuGlnLysGlnLysValThrCysLeuSerGlnLysPhe 971
 Db 3424 ACCTTGCAGTAAACA----- 3441
 QY 972 LeuProAsnGlnValGlnLysLysAsnGlnLysSerLys----- 985
 Db 3442 -----AACCAATATAGATTGCTCCAGCAGAAATCAGTGAAGAAACCAAGCAATTATCCAG 3495
 QY 986 -----AlaSerGlnGlnLysMetLeuLysGln 995
 Db 3496 AAGTTAATCACAAGTAAACAGGATGCAAGTGAAGGAGCTCCGTAAGCACTTGAAGGAA 3555
 QY 996 GluGlnGlnHisSerAlaSerLleLysSerLysGlnLysLleLysMetGlnGlnLys 1015
 Db 3556 ACACTGTGATTAAGTCCACCTTGACAGTAGTAGTGAAGCACTGGAACCAAGAACTAGAA 3615
 QY 1016 GluGlnLleLeuGlnLysThrAspGlnValThrHisThrGlnSerLysValGln----- 1033
 Db 3616 GMAAGATAGTGGCTGTGAAAGAAAGAGAGCAACTTCAAAAGAGTACAGAGAAAGCC 3675
 QY 1034 -----GlnThrGlnGlnGlnLysLeu----- 1040
 Db 3676 TTAACTCCCGCAAGCAATTTCTTAAAGCAACAGAGAGAAAGAAAGCACTTCAGGAG 3735
 QY 1041 GluMetLysLysMetHisAspAspLeuPheGlnLysTyrLleArgAsnLysSerGlnAla 1060
 Db 3736 GACTTAAGCAACAGAAAGATGAC-----TATTAATGCTTGAAGAAAGCTT 3783
 QY 1061 GluAspLeuLysArgGlnMetGlnLysAsnLeuLysGlnLysMetLysSerValGlnValLys 1080
 Db 3784 GATGAGCAAAAGCAAGAAATGAAATATTGAGACAGTGAAGCACTCCAGATTCAA 3843
 QY 1081 IleAlaAspThrLysHisGlnLysGlnGlnLysThrLleArgAspLysGlnLysLeuHis 1100
 Db 3844 GTAAGGAATCC----- 3855
 QY 1101 GluLysLysTyrPhePheGlnAlaMetGlnThrLlePheProLleThrProLeuSerAsp 1120
 Db 3856 -----ATAGACGGA 3864
 QY 1121 SerLeuProSerLysLeuValGlnGlnLysAsnSerGlnAspProLleGlnLysAsn 1140
 Db 3865 AAACCTCCCAAGCAGACACAGCAGGAGATGCTGTTCTCCACTCA-----GGT 3912
 QY 1141 TyrHisAsnLeuLleAlaLeuAlaThrGlnArgAsnAsnLleMetValCysLeuGlnThr 1160
 Db 3913 TTAGAGAACTTTATTCAAGCCACAGAAAGCAACATCAGCAACCTGTTTAAGAGTCC 3972
 QY 1161 -----GluArgAsnSerLeuLys-----Glu 1167
 Db 3973 AACTTGTGCCAGACTGGCTTCTCATTTCTGAAGAGTCCAGCTGCGTCCAGGGCGGAACT 4032
 QY 1168 GlnValLleAspLeuAsnThrGlnLeuGlnLysSerGlnAlaGlnSerLleGln----- 1185
 Db 4033 TCTGTTGCCACATTAAGGCCACAGCTGAAGAAATTAAGAGCTGAAGAAATTAAGTGAAG 4092
 QY 1186 -----LysSerAspLeuGlnLysProLysGlnAsp-----Leu 1196
 Db 4093 TTGAAGTATGATTTCAACAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4152
 QY 1197 GluGlnGlnValLysLeuLeuLeuGlnMetGlnLeuLysGlnLysLysLeuThrAsp 1216

```

Db 4153 CAACGACGATTAATTAACAGAGGTTTAGAAATCGAGACTTAAGACAGTATCCATGAA 4212
Qy 1217 SerLInleuSerLleuLysLeuGlnLeuGlnLeuGlnLeuGlnLeuGlnLeuGln 1236
Db 4213 GCTGAAGTCCATGCCGAAAGGCTGCAG-----CAAGAAATTGGAA 4251
Qy 1237 ThrLeuGlnGlnLeuMetLysAsnIleThrIleGlnArgAsnGlnLeuGlnThrAsnIle 1256
Db 4252 AGCAGCCACACACAAATGCTGCGCTAGACACTTAAAG---GAATTCCAACTTAAGTGTG 4308
Qy 1257 GluAspLeuLysAlaGlnHisAspSerLysGlnAspLeuSerGlnAsnIleGln 1276
Db 4309 GATGAAGTGCAGAAATCATACAGCAAAAGAGAGACGTTAGACTCTTGTGACAA 4368
Qy 1277 SerIleGlnThrGlnAspGlnLeuArgAlaIleGlnGlnLeuArgGlnLeuGln 1296
Db 4369 CTGTGTGAGAAAGACAGCTCTCATTAATACAGACAGAGATTAATGAACAAAGAAAT 4428
Qy 1297 LeuValAspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIleSerSerProAsn 1316
Db 4429 TTATTAAGGCTCGCATACACAGCTA-----GAATGCCAAGCCAAAGAG 4473
Qy 1317 HisAspAlaValAlaAsnGlnGlnLysValSerLeuGlnValAsnSerLeuGlnSer 1336
Db 4474 CATGATGAGAGATTAAGACAGCTACAGGTGAACTTTGTGAATGAACAAACAGAA 4533
Qy 1337 GluMetLeuArgGlnValArgAspGlnLeuGlnThrSerCysLys-----AlaLeu 1353
Db 4534 GAGATTGAGAGAAAGAAAGTAGCAAGCAACAAATACAAAGAAACCTGCAAGCTGCCCT 4593
Qy 1354 ValSerGlnLeuGlnLeuLeuArgAlaHisValLysSerValGlnGlnGln----- 1370
Db 4594 ATTTCCGAAAGAAAGAACACTTAAGAAAC---AAAGTCTCCAAAGAGAAATTGCTTTG 4650
Qy 1371 -----AsnLeuGln-----IleThrLysLysLeuAsnGlnLeuGlnLysGlnLeu 1386
Db 4651 GCCAGAGTACCATTTGAACGCTCAGCAAGCTCTGCGCAGATGTGAAAGCAAGTTCT 4710
Qy 1387 GlnLysSerGlnGlnGlnValLeuLysSerMetLeuGlnAsnLeuLysGlnAspAsn 1406
Db 4711 GCTGAATTAATTAAGAAAGATACGCTTTAGAAAGTTAGCTTTCTCAAGAAAGAA 4770
Qy 1407 AsnLys-----LeuLysGlnGlnAlaGlnGlnLysSer 1417
Db 4771 GACAAACTCATTAAGAAATGACAGAGCTTTATTGAAATTCAGAGCTCAGCAGCTCC 4830
Qy 1418 SerLysGlnAsnGlnPheSerLeuGlnValPheSerGlnSerGlnLysLeuValAsp 1437
Db 4831 TGTGAAGTCTAAAACTAGCTCTAGAGGCTCTACTGAAAGCAAGAAAGTTAGTGAA 4890
Qy 1438 GlnIleGlnValLeuLysAlaGlnLeuLysAlaIleGln-----GlnArg 1452
Db 4891 GAAATTAATCTTTGAAATCTTCT---AAGATTGCAGAAAGTACTGAGTGCAGAGAA 4947
Qy 1453 LeuGlnIleLysAspArgAspTyrPheGlnLeuValGlnThrAlaAsnThr----- 1469
Db 4948 CACAAGAGAGCTACAAAAGAGATATGAATCTCTGCGAGCTCTTGAATGTTAGTAA 5007
Qy 1470 -----AsnLeuValGln----- 1473
Db 5008 GAACGACAAAGATTCAGCATGTGTGAGAGCTGTGAGGCAAGAAACAAGAACTGTAT 5067
Qy 1474 GlnLysLeu-----GlnThrProLeuGlnAlaAspHisGln 1485
Db 5068 GGCAGATTAAAGACACAGAGGCAACAGAAAGAGACAGAAAGAGCTTCAGAGAACT 5127
Qy 1486 GluAspSerIleAspArgArgSerGlnGlnMet----- 1496
Db 5128 GAGCAGAGAAATGAGAAATGAAGAAAGATGAGAAAGTTGCTTAATCTAAACAGCAG 5187
Qy 1497 -----GlnIleLysValLeuGln 1502

```

```

Db 5188 AAAATCTAGAGCTGGAAGAGAAATGACCGGCTTAGGCGCAGAGAGTCCAGCAG 5247
Qy 1503 GlnLys-----LeuGlnArgAsnGlnTyrLeuLeuGlnArg 1514
Db 5248 GATACAGCTAAAGAGCTGTATGGAACACTTCTTTCTCCAAATGCCAGCATGAGAAAGAA 5307
Qy 1515 LeuGlnGlnGlnLysLeuGlnLeuSerAsnLysLeuGlnIleGlnLeuGlnMetGln 1534
Db 5308 CTGGAAGAGCTCAAAATGAG-----TATGAACCCCTTTCTAAGAACTTTGAG 5355
Qy 1535 ThrSerValLeuLeuLysAspAspLeuGlnGlnLysLeuGlnSerLeu-----Leu 1551
Db 5356 TCTTAATGTGTGAAGAAAGCTCTTAAGTGAAGAGCTTCAAGATTTAAGCATGAGATA 5415
Qy 1552 SerGlnAsnIleIleLeuLysGlnAsnIleAspThrIleThrLysGlnHisSerAspThr 1571
Db 5416 GACGTAATGTATCTTAACAAAGCTTAACCTAGAGGCCAGCAAGAAATGATTAACCAAG 5475
Qy 1572 GlnAlaGlnLeuGlnLysThrGlnGln-----GlnLeuGlnLeuAlaLysAsnLeu 1588
Db 5476 AATGCACATGAAGAGGAAACACAGCTGTATACAGGTGAGAGCTGAGAAAGCAAGACTCTG 5535
Qy 1589 AlaIleAlaIleSerAspAsnCysProIleThrGlnGlnLysGlnThrSerAlaAspCys 1608
Db 5536 AGTATGACACAGAAAGCTTACATGT-----TCAGAAATGCGTTTCATCAGCCAAAGAGT 5586
Qy 1609 ValHisProLeuGlnGlnLysIleLeuLeuLeuThrGlnGlnGlnLeuHisGlnLysThrAsn 1628
Db 5587 GCCAACCTGCTGTAAAGATTAAGATTCAGCTCAGTGAATGAATTAAT----- 5634
Qy 1629 GlnGlnGlnLysLeuLeuHisGlnLysAsnGlnGlnGlnAlaGlnValGlnLeuLys 1648
Db 5635 -----AATCACTCAAGCAGAGATTGAT---CAGCTCAA 5664
Qy 1649 CysGlnValAlaGlnHisLysMetLysSerMetIleGlnSerLysSerSerLeuGlnSerLeu 1668
Db 5665 GAAAGAAATGCTGATTTAGAGAGGAGAAAGCAAGAAACAAAGAAATTTAGCCAGACTTAA 5724
Qy 1669 GlnHisGlnLysHisAspThrGlnGlnGlnLeuLeuAlaLysGlnGlnMetGlnVal 1688
Db 5725 GAAATGAGAAATAATCCTTACTGAGTCAAGATTCACAAAGATGTGAATTAAGATG 5784
Qy 1689 ValThrGlnGlnLysLysLeuLeu-----GlnGlnThrHisGlnHisLeuThr 1704
Db 5785 CTTCAGAGAGAAATGAACCAAAATGAATGAACTGTTAATTCAGCAAAATCCAAAGAAACTC--- 5841
Qy 1705 AlaGlnValAspHisLysLeuGlnAsnIleGlnLeuGlnLeuAsnPheLysAsnGlnAla 1724
Db 5842 TCCAGAGTTACCAAACTTAAGAGAGACAGCAGAA----- 5874
Qy 1725 GlnGlnLysThrThrLysGlnGlnCysLeuLeuAsnGlnValAsnLysGlnLeuGlnGlnSer 1744
Db 5875 GAGAGAAAGAGATGATTGGAAGAGGCTTATGAATTCATTAATGCAAGAACTTAATGGAAGC 5934
Qy 1745 GlnHisArgLeuGlnCysGlnIleGlnLysLeuMetLysSerLeuLysAspLysGlnSer 1764
Db 5935 -----ATTGCGAATTAAGTCTCAGAGATGTTAACAGATGCCCAATA 5973
Qy 1765 AlaLeuGlnThrLysLysGlnSerGlnGlnLysValIleAsnLeuAsnGlnGlnMetGln 1784
Db 5974 AAAATAGAGCTTAATG---GAATCTGAAATGAAG---AACCTTAAGAAAG----- 6015
Qy 1785 MetValMetLeuGlnMetGlnLeuLysAsnSerGlnIleArgThrValIleAlaGlnArg 1804
Db 6016 -----TGTGTAGTGAATTTGGAAGAGAAAGAGACAGACTTATGTCAGAGAA 6063
Qy 1805 AspGlnLeuGlnAspAspLeuArgGlnSerValGlnMetSerIleGln----- 1820
Db 6064 ACTAAGTGGAAATCAGAAATGCAAGAGAAATATTTGCAAGAAATATCAAGAGTGCAGAA 6123
Qy 1821 -----ThrGlnAspAspLeuArgGlnAlaGlnGlnValAlaLeuGlnGlnGlnLysAsp 1837
Db 6124 GAACCCGAAATTAAGAAAGCCATGCAAGAGAACTTCAAGAACTGTTAAGAAAGAAACAA 6183

```

QY 1838 LysValGlnGluLeuThrSerGlnIleSerValLeuGlnGluIysIleSerLeuGlu 1857
 Db 6184 GAAGTAAAGAGCTACAGAGAGATGTCATCAGTATCAAGAGAAAAATAGGCTCGAG 6243
 QY 1858 AsnGlnMetLeuTyrAsnValAlaThrValIys-----GluThrLeuSer 1872
 Db 6244 AGA-----ACTGTTAAAGCTTAGAATTTGTTCAACAGTATCT 6282
 QY 1873 GluArgAspAspLeuAsnGlnSerIysGlnIleSerPheSerGlnIleGlu----- 1889
 Db 6283 CAAAAA---GATTGGAAATTAACCAAGAAAAAATCTGGCTCAAGAGCTGAACACCGCAAA 6339
 QY 1889 ----- 1889
 Db 6340 AAGCACAAGCAGATTAGCTAGCTCAAGTCTCTAGATGACACTCAAAAGTGAGCA 6399
 QY 1890 -----ThrLeuSerLeuSerLeuIysGlnIysGlnPheAlaLeuGlnAlaGluIys 1907
 Db 6400 GCAGAGGCTCTAGCAGACATCTCAAGTGAAGAAAGACTGACATCAATTAAGATCA 6459
 QY 1908 AspLysAlaAspAlaAlaArgIysThrIleAspIleThrGluIysIleSerAsnIleGlu 1927
 Db 6460 GTTAAAGCCAGATGAACAAAGAGATGAAGATCTTGAGCGAAGACAGCAAGCGCGAA 6519
 QY 1928 GlnGlnIleuGlnGlnAlaThrAsnLeuIysGluThrLeuIysGluArgGluSerLeu 1947
 Db 6520 GAGAGACACTGAAAGAGAGAGATATGCAAGAGAAACGT-----GATGCTTTG 6570
 QY 1948 IleGlnCysIysGlnIleuAlaLeuAsnThrGlnIleAsnArgGluThrLeuIysSer 1967
 Db 6571 CGCAGAGAAAGATGCCACTTGGAAGACAAATGAGAGATTCAGATTCTTGACAAAG 6630
 QY 1968 LysAspLeuAlaLeuGluIysMetGluGlnGlnIysArgAspGluAlaAlaAsnIysValIle 1987
 Db 6631 AAGACAGAGAGATTCAGCAACTTCAGGAAAACTTGACAGCTGCTGACCGACTTGA 6690
 QY 1988 AlaLeuThrGlnIysMetSerSerLeuGlnGlnIleAsnGlnAsnValThrThrLeu 2007
 Db 6691 GCCTTACTAGAGCATGCTCTCCCTCCAGATGATGCTGACAGGATGATGATGAGCT 6750
 QY 2008 LysGluGluIys-----GlyIysGluThrPheTyrLeu 2019
 Db 6751 AAGAAAGCGAGAGAGATTAGTATGATGATCAACAAAGAAAGAAATTTAGATC 6810
 QY 2020 Gln-----ArgProSerIysGlnIleAsnSerSerSerGlnMetGlu 2032
 Db 6811 AAGAGATATATTCAGATGTTCTTAAGATCAACTTACACATGTCATTCATATGAAA 6870
 QY 2033 GluLeuArgGluSerLeu----- 2038
 Db 6871 GAATTTAAAGATTAACATTTCCAGGCTTGAAATGACATGACAGATTTGGAGTCCAGGCC 6930
 QY 2039 -----LysThrIysAsp 2042
 Db 6931 CAGACAGAGGTCACGCTCAGCAGAGGCTGTGATCTCTACAGGGGAGAAACAAAGAA 6990
 QY 2043 Leu-----GlnLeuGlnGlnAlaGluIysGlnIleSerGlnAlaThrAsnGluIys 2060
 Db 6991 CTTTTCCTCCAGCTAGAGAGACAGCCACCTATACACAGTCTCGAATGATTAAGCT 7050
 QY 2061 AsnLeuThrAlaLysIleSerSerLeuGlnGlnIle-----LeuGlnAsnAla--- 2077
 Db 7051 AAGTGGATACAGACTTAAGAGTCTCAAGACCGATTGATTAAGTAACTCTTTA 7110
 QY 2078 -----SerIleLeuAsnGlnAlaValSerGlu 2086
 Db 7111 GAAAAATGTAGCAAAAGAAAGAACTTGAAGGATCATTAAGGCGCAAGAGCGCTGAT 7170
 QY 2087 ArgGlnAsnLeuArgHisSerIysGlnGlnIleValSerGlnLeuGlnLeuSerLeu 2106
 Db 7171 ATTCAAAATTTCTAGTTCACTTATGAACACTGAGACATGATCTT---CAGGCCCTCAGA 7227

QY 2107 ThrLeuLysSerArgAspHisAlaPheAlaGlnSerIysArgGlu----- 2121
 Db 7228 GAAGTACACAGTAGGCTGATGAGAAATTAATATGAAGAGCAAAAGATTTAAGCTCG 7287
 QY 2122 -----LysAspGlnAlaValAsn---LysIleAlaSerLeuAlaGlu----- 2134
 Db 7288 CTTTCTGGCAGAGAGAGCATTCACAGATAGCTATTTGCTGAACCTGCTGACCAATGAT 7347
 QY 2135 -----GluIleLys-----IleLeuThrIysGlnMetAspGlu-----Phe 2146
 Db 7348 AAGAAATTAAGAGCTGTAAGAAACCTGCTGCTCCAGAGGAAAGAGAGAAATTTGTTTA 7407
 QY 2147 ArgAspSerIysGlnSerLeuGlnGlnIleSerSerHisLeuSerGlnLeuCysThr 2166
 Db 7408 GAAGAGAGAAACAAAGCGCTTGATTAACCAATCCACTTATGTAACCTGAAACCTGAAAAC 7467
 QY 2167 TyrIysThrGlnLeuGlnMetLeuLysGlnGlnIysGlnAspIleAsnAsnIysLeuAla 2186
 Db 7468 ATCAAA-----AAGGAAACATTCACGCAAG----- 7494
 QY 2187 GluLysValIysGlnValAspGluLeuGlnIleSerSerLeuIysGlnLeu 2206
 Db 7495 -----GCACAGTGGATTCCTTTGTTAAATCCATGCTCTCTCCAAATGATGCA 7545
 QY 2207 AspGlnIleGlnMetGluLeuArgAsnGlnIysLeuArgAsnTyrGluLeuCysGluIys 2226
 Db 7546 GACCGCTAGTGGGTGACATCAACAGCTGGAAGAGCCATCTCTCTATTA----- 7596
 QY 2227 MetAspIleMetGluIysGlu-----IleSerValIleAsnArgMetGlnAsnGluPro 2244
 Db 7597 -----ATCTTGGAAAAAGCAACACTCAACAGGCTGCTGCGAGAAATTAACCTT 7650
 QY 2245 GlnGlnGlnIysGlnAspValAlaGluArgMetAspIleLeuGlnSerArgAsnGlnGlu 2264
 Db 7651 AAGAAAGAAATACAGAGGCTTGAGAAATGATATGATGATCAATTCGAGAAATCCCAAG 7710
 QY 2265 Ile-----GlnGluLeuMetGluIysIleSerAlaValIysSer 2277
 Db 7711 CTAGATGCAAGACTGATCATATAGAGAGACCTGAAACCACTGTATTAACAATAAGAGAC 7770
 QY 2278 GlnGlnIleThrLeuSerSerLeuSerSerGlnLeuGlnIysGlnThrGlnAlaHis 2297
 Db 7771 AGCCAAACAAAGACGCTCTTCTTGAAGTTCACTTCAACGAAATTAAGAGCTGGA-- 7827
 QY 2298 LysHisCysMetLeuAsnIleLys-----GluSerLeuSerSerThrLeuSer 2313
 Db 7828 AATATATCTTAATTAAGAGAAAGCTGAAGAAATCTGAGAAAGAAATGAGATCTGCGG 7887
 QY 2314 ArgSerPheGlySerLeuGlnThrGlnHisValIysLeuAsnThrGlnLeuGlnThrLeu 2333
 Db 7888 AGTCTCTTATGCTCTTCAAGAGAGAAACAAAGATTTATCTTAAGAGATTTGAGAGCTTTG 7947
 QY 2334 ---LeuAsnIysPheLysValIleTyrArgThrAlaAlaValIysGluAsp----- 2349
 Db 7948 AAGATATCTATATCCAGCTTACACAGACATGACAGCTTGCAGAAAGAGTACTTTA 8007
 QY 2350 -----HisSerLeuIleLysAspTyrGlnLysAspLeu----- 2360
 Db 8008 GCACTCTATCATGCCAGCTTAAAGTAAGAAAGAGAGGCTACACAGGTTAAGTCTTTG 8067
 QY 2361 ---AlaAlaGlnIysArgHisAspGlnLeuArgLeuGlnIysLeuGlnIys 2379
 Db 8068 TTTTCTCTCTCCCAAAAGAGATTCAGACTGGAAGAGAAAGATTTGTTCTGTTCAAAAG 8127
 QY 2380 HisGlyArgIysTyrSerSerAspSerAlaSerGlnLeuLysPheCysGlnIleGluPhe 2399
 Db 8128 GAAGCTCCCAAGAGTACAGTAATTAAGATTAACAGAG--AAGAAATTAAGCAT 8184
 QY 2400 LeuAsnGlnLeuLeuPheLysIysAlaAsnIleLeuGlnValGlnAspPheSer 2419
 Db 8185 CTTGATCATGATGACAGGATTAAGAAATGAAGTGAACGAGAGAGAGAGTGGCA 8244
 QY 2420 GluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlnIleHisLys 2439

```

Db 8645 GAGGACAGACAGATTTGGTGAGATGGAGAACAGAAATTACATGCTGACCAAGAAAT 8304
OY 2440 LysGlyPheMetGlnThrPheGlnGluPheGly-----AspLeuHisValAsp 2455.
Db 8305 AAAGGTCTCAGACAGCAAAATTCAGTCTTTGGAGGTCTATAGTTCCTTCAAAATAGT 8364
OY 2456 AlalysLeuSerGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2475
Db 8365 AGAGATCATGCGATGAGAGAACTGATGAAAGAAATATGATGCGACGCTGAG 8424
OY 2476 LeuLeuThrLysArgLeuValAlaValAlaGlnSerLysIleGlnArgGluIleThrVal 2495
Db 8425 GAATTGGCA-----CAGTTGAAAGAA-----CAGGACTCTTAAACAGAGAGAGATGCT 8475
OY 2496 TyrLeuAsnGlnPheGlnAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2515
Db 8476 CTCTCTTCTGAAACCGCTTTTCATGAACTCCACTGAGGAAATAGCTTCTCACCTT 8535
OY 2516 ArgArgMetGlnHisGlyProSerAlaSerValMetGlnGlnGlnAlaArgLeu 2535
Db 8536 GAGAACTTACCAACAG-----CTCTATCCAAAGATGAGCAATTG 8577
OY 2536 LeuGlyIleLeuLysThrValGlnAspLysLysLysLysLysLysLysLysLysLysLys 2554
Db 8578 CTTCACCTTGCTCCACAACTAGAGATTTCTATACCAAGTCAGTCTTTCCAGGCT 8637
OY 2555 -----MetLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2567
Db 8638 ATGGCCAGCTCTCGAATGAGAGATGACCTGTAATGATGAGTGGAGAAATTCGAAAG 8697
OY 2568 MetHisLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2587
Db 8698 TCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8745
OY 2588 AlaGlnLeuAsnAlaMetGlnValLysLeuThrLysLysGlnAspAsnLeuGlnAla 2607
Db 8746 GCTGAAGTACAGAGATTTAAAAAAGCTATGCTCCTCCTCAATGACAGAGAGACTA 8805
OY 2608 MetLysGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2627
Db 8806 CTGAAGAAATTTGAAATCTGCAGCAGCAATAC-----TTCAGATTAACTCA 8853
OY 2628 GluIleAspAsnLeu-----LysThrLys 2635
Db 8854 GAGATCAGTGAATTCATCCACTGAGGCTCACTGAGGATTCAGATCAAGATCAACAAA 8913
OY 2636 ValValLysIleGlnMetGlnLysIleLysTyrSerLysAlaThrAspGln--GluIle 2654
Db 8914 GCATTTCAGATTATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8973
OY 2655 AlaTyrLeuLys----- 2658
Db 8974 CATCAGCTCAGAGATGAGAAAGATTCCTGGGAAATACATGAGAGAGAGAGAGAGAGAGAG 9033
OY 2659 -----SerCysLeuGlnAspLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 2676
Db 9034 TACCTTATGGCTATCTCAGATTAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9093
OY 2677 ArgAlaGlnAlaAspAsnAspThrThrValCysValProLysAspTyrGlnLysAlaSer 2696
Db 9094 GAATTGGGTCTTCTCTCCACAGACTCCTCAAAAGTCAATACCAAGAGAGAGAGAGAGAG 9153
OY 2697 ThrPheProValThrCysGly-----GlyLysSerGlyIleValGlnSerThrAla 2713
Db 9154 TCC-----CCAGAGACATAGCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9210
OY 2714 MetLeuValLeuGlnSerGlnLysAlaAlaLeuGlnArgGlnLeuSerHisTyrLysLys 2733
Db 9211 CTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9264
OY 2734 LysTyrHisHisLeuSerArgThrMetSerSerSerLysAspArgLysLysThr----- 2751

```

```

Db 9265 AGAATTACAGCAACTGACAGCAACTTCTCTCAGTACTGGAAGAGAGAGAGAGAGAGAGAG 9324
OY 2752 ---LysAlaLysSerAspAlaHisSerHisThrGlySerHisArgLysSer--- 2769
Db 9325 ATTCAGCTCTCGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9384
OY 2770 -----ProHisLysThrGlnThrLysThrGlnHisGlyProValThr 2782
Db 9385 TTAATCAGTGTGCGAGTCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9444
OY 2783 ProGlnArgSerGln---MetProSerLeuHisLeuGlySerProLysLysSerGlnSer 2801
Db 9445 ATAGATGTTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9504
OY 2802 SerThrLysArgValValSerProAsnArgSerGlnLysThrSerGlnLeuValMetSer 2821
Db 9505 GAGGAAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9564
OY 2822 ProGlyLysThrGlyMetHisLysHisLeuSerProSerLys----- 2836
Db 9565 AGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9624
OY 2837 -----ValGlyLeuHisLysLysArgAlaLeuSerProAsnArgSerGlnMet 2852
Db 9625 GCTGAGAAATGCTCTCTCTGTGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9684
OY 2853 ProThrGlnHisValIleSerProGlyLysThrGlyLeuHisLysAsnLeuThrGlnSer 2872
Db 9685 GACCTCTCCGAGAGCTCTATCATTTGGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9735
OY 2873 ThrLeuPheAsnLeuSerSerProCysArgGlnGlnLys 2886
Db 9736 CTGTTAATAGATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9777

RESULT 4
US-09-960-253-163
Sequence 163, Application US/09960253
Patent No. US20020123619A1
GENERAL INFORMATION:
APPLICANT: Benson, Darin R.
APPLICANT: Monamath, Raedoh
APPLICANT: Lodes, Michael J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.556
CURRENT FILING DATE: 2001-09-20
CURRENT APPLICATION NUMBER: US/09/960,253
NUMBER OF SEQ ID NOS: 187
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 163
LENGTH: 10096
TYPE: DNA
ORGANISM: Homo sapiens
US-09-960-253-163

Alignment Scores:
Pred. No.: 1,15e-64 Length: 10096
Score: 1152.00 Matches: 740
Percent Similarity: 37.41% Conservative: 563
Best Local Similarity: 21.25% Mismatches: 1160
Query Match: 7.80% Indels: 1022
DB: Gaps: 140

US-09-150-867-1 (1-2954) x US-09-960-253-163 (1-10096)
OY 81 PheAlaTyrGlnGlnThrSerSerGlyLysThrThrMetMetGlyThrProAsnSer 100
Db 465 TTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
OY 101 LeuGlyIleIleProGlnAlaIleGlnGlnValPheLysIleIleGlnGlnIleProAsn 120
Db 495 -----AAACAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 530

```


QY 121 ArgGluPheLeuLeuArgValSerTyrMetGluIleTyrAsnGluThrValLysAspLeu 140
 Db 531 TGTAAATTCGAGCTTGAAGAAAGCCAAAGCTGCGCAGTGCAGATGCTCTCTGAAAT 590
 QY 141 LeuGlyAspAspArgArgLys-----LysProLeuGluIleArgGluLysPheAsn 157
 Db 591 CCATGCAGATACACCAAAAAATTTTACACTCCCTACACCAAGCAATATATATAGT 650
 QY 158 ArgAsnValTyrValAlaAspLeuThrGluGluLeuValMetValProGluHisValIle 177
 Db 651 GGTTCCAGATAT---GAGATCTTAAAGAAAAATATATATATATATATATATATATAT 698
 QY 178 GluThrPheLysGluGluLysAsnArgHisTyrGluThrLysMetAsnAspHis 197
 Db 699 -----GACGAAAAAGATTAGAGCAGAGAGGTTTAAACCTTCGAGGCT 740
 QY 198 SerSerArgSerHisThrIlePheArgMetIleValGluSerArgAsp-----ArgAsn 215
 Db 741 AAAAAGCAACCCAGACTCTCCACAGCCACATGATATCCGCGCATTCGCCGCGCAT 800
 QY 216 AspProThrAsnSerGlu---AsnGlyAspGlyAlaValMetValSerHisLeuAsnLeu 234
 Db 801 CAGGCTTATATCTGCTGCTTCTCAGCAGCAGAGAGAAAGCCCAAGTATCTTCTCA--- 857
 QY 235 ValAspLeuAlaGlySerGluArgAlaSerGluThrGluAlaGluGluValArgLeuLys 254
 Db 858 -----TCATATTTCTCAAGAAACT----- 875
 QY 255 GluThrLysAsnIleAsnArgSerLeuPheIleLeuGluValIleLysLysLeuSer 274
 Db 876 -----CCAAATTAGAGAGATTCTCT 896
 QY 275 AspGluGlnAlaGlyPheIleAsnTyrArgAspSerLysLeuThr----- 290
 Db 897 GCATCTTACTTTCTGGGAACTAGAGTGACTCCCAATCTCATCATCTTCCAAATAGCG 956
 QY 291 ---ArgIleLeuGlnAsnSerLeuGluGlyAsnAlaLysThrValIleLysThrIle 309
 Db 957 AAAAGAGATCTAATAGCAGTTCTTGGCAATTCTACG----- 995
 QY 310 ThrProValSerPheAspGluThrLeuSerThrIleGlnPheAlaSerThrAlaLysHis 329
 Db 996 AGCCCTCATCTTTGGAT-----CAATTAAAGCCGCAAGATCAAG 1037
 QY 330 ValArgAsnThrProHisValAsnGluValLeuAspAspGluAlaLeuLeuLysArgTyr 349
 Db 1038 CTAAGAAAC-----AAGATTATAGATTG-----GAACTACGCTGCAAGACAT 1082
 QY 350 ArgGlyGluIle-----LeuAspLeuLysLysGluLeuLys--- 361
 Db 1083 GAAAAAGAAATGAAGGCCAAAGTGAATTAAGTTCAAGAACTCCCACTCACTGGAGAAA 1142
 QY 362 ---AsnLeuGluSerSerSerGluThrLysAlaGlnAlaMetAlaLysGluGluHisThr 380
 Db 1143 GCAAAAGTGGAAATTAATGAAAAAGAGAAAGTTTGAAACAAATGATGGATGAACAGT 1202
 QY 381 GlnLeuLeuAlaGluIleLysGlnLeu-----HisLysGluArgGluAspArgIle 397
 Db 1203 AGAACACACGACATAGCAGCCAGCGTCAACCAAGATATCTCATTTGCAACAAAAGT 1262
 QY 398 ThrHisLeuThrAsnIleValAlaLysSerSerGlnLysGluAsn----- 413
 Db 1263 AAAAAATTCGAGAGATAGTTGATGCTCAGCGACACAAATATGCAAAATGCGCAATGTCT 1322
 QY 414 ---AspGluArgValArgLysArgValArgValThrThrAlaPro---GlyLysIleGln 431
 Db 1323 CTGGAAACGAAATTAAGAAAAAGAAAGAGTTTCAAGAGAGCTCTCCGCAACAG 1382
 QY 432 AsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeu---ProGluLys 451
 Db 1383 CGTCTCTTCCAAACACTGCAGCAGAGATGCATCCAGATGAAGGCCACAGATCAACCCAGAG 1442
 QY 451 nHeserLysLysAlaLysPheSerAspMetProSerPheProGluIleAspAspSerVal 471

Db 1443 TTAACAGCAG-----CCAAAGAT 1460
 QY 471 LysThrGluPheSerAspPheAspAlaLeuSerMetLeuAspSerArgLysIleAs 491
 Db 1461 ATGCACACAGCTCTGAGGCTAGACTGATTAACATCATCATGATTAACCAACACTGAA 1520
 QY 491 PaaGluThrAsnLeuAlaSerLysValThrHisArgGluLysThrSerLeuHisGlnSe 511
 Db 1521 AACAAATTCGAGAGTTTAAGCAAAAGTTGTCG-AGAGTGAACAGAGGCTTCCAGCGCG 1579
 QY 511 rMetLeaspPheGluGlnLysSerAspSerValGlnPheHisAspSerSerLysGluAs 531
 Db 1580 TCAGATCAAGAGATAGCTGAGAGAGAGCAGTGGAG-----GAAATGACAGAGAAA 1633
 QY 531 nGluLeuGlnThrLeuProLysAspSerGlyAspMetAlaGluLysArgLysAlaSerP 551
 Db 1634 CAACCTCTTAAGAGTACCTGAGCAAAAGCCACAGAAAGTCTGCCACTGAGAGCAGA 1693
 QY 551 eGluLysGluIleThrSerLeuGlnGlnGlnLeuGlnSerLysGluGluLysGlu 571
 Db 1694 ACTCAAGAACATCAAAAGAGTGTAAATCAGAGCCAGAAATTTGCAAGAAATGAAAC 1753
 QY 571 uLeuValGlnSerPheGluLeuLysIleAlaGluLeuGluGluGluLeuSerValLysAl 591
 Db 1754 GAAAGATACCTCTCAGGAAACCATGTTAAGAGATCTTCAAGAAATAAATA----- 1802
 QY 591 aLysAsnLeuGluMetValThrAsnSerArgGluHisSerIle----- 605
 Db 1803 -----AATCAGCAAGAAAACTCTTGCCTTACAAAACACTGAA 1840
 QY 606 -----AsnAlaGluValGlnThrAspValGluLysValAlaArgLysGlu 621
 Db 1841 GCTTCTGCTGCTGATCTGGAAGAAACAGCAGATGTTCTCAAGACCTTTGAAAGAAAG 1900
 QY 621 uMetSerValLeuGluLysAspSerGlyTyrAsnAlaSerAsnSerPheLysGluAspSer 641
 Db 1901 A-----GAAATCATCATTTGAACCACTTAATATATATATATATATATATATAT 1945
 QY 641 rValAspGlyLysArgLeuSerSerSerHisAspLysLysGluLysArgLysMetLe 661
 Db 1946 GAAAGAGTCCAAAGCCTTCTGAGTGT-----TT 1975
 QY 661 uGluGlnLysIleValAspLeuGluIlePheIleGluAsnLeuAsnLys----- 677
 Db 1976 AGAGTTAAAAAAGCAAGATATGAGAAATTTGAAGAGAGAAAGAACTGTTTCTTCTTG 2035
 QY 678 -LysSerGluAsnAspLys-----GlnLysSerSerGluGluAspPheMetGluSe 694
 Db 2036 GAAAGAGTGAAGCAAAAGCTTTTAATCTAGATGAAATGAAAGAAAGAACTTCAAGAG 2095
 QY 694 r-----IleGlnLeuGly----- 698
 Db 2096 TAAATTAATACCTTGAACCTTGTGAAAGCAGCAGCAATTAATTAATGATGAATACAA 2155
 QY 699 -----GluAlaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeu 715
 Db 2156 CGAGAGATGAAGAGCTGAGATGACAGAGAAACCTTAAGTGTGAGATGAAGAACT 2215
 QY 715 tArgAspAsnPheAspAsnIleIleLeuGluAsnGluThrLeuLysArgGluIleLys 735
 Db 2216 TCACAACTGTTACACAGTAACTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2275
 QY 735 pLeuGluArgSer----- 739
 Db 2276 GCTTACAGCAAGAAAGCTGATCTTCAATCAGAAACATCAAGAGAAATAGAAATATGTC 2335
 QY 740 -LeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGluLysGluThrGlnLysGlu 759
 Db 2336 TTGGAAGACTTCTCACTGAGGCAAGTTGAAGATCA-----GA 2377
 QY 759 uHisGluAlaGluLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsnAlaGlu 779

Db 2378 ACACAGCTTCAGTTACTGTCA-----AATGAATATGACAAAGACCG 2422
 QY 779 UMETTYASnglnAsnleuglnuAspLeuGlnThrlYsThrlYsleuLysGlnG 799
 Db 2423 GTGTAC---CAAGACTTGATGCCGAAATATGAGAGCCCTCAGGATCTGTAAATCCAA 2479
 QY 799 nglnlGlnleu----- 803
 Db 2480 AGATGCTCTCTGTGTGACAAATGAAAGATCATCAAGAAAGCTTTGGCTTTGATCAGCA 2539
 QY 803 ----- 803
 Db 2540 GCCTGCATGCATCATCTCTTGGCAAAATATATATGAGAACAGAGCATGCCCTCAGA 2599
 QY 804 ----AlGlnleuArgLysArgAlaAspAsn-----LeuGlnLys 815
 Db 2600 GAGAGAGGAATGTGCTGTAGAGAGACCAAGATCCGAAAAATTCGCCATCTCACAAA 2659
 QY 815 sLysValArgAsnPhesAspLeuSerValSerMet-----GlyAspSerG 830
 Db 2660 TAGAGTGTATCTACTGTGAATTTTTCATTAGAGTCTCAAAAACAGATGAAGCTCAGACCTGCA 2719
 QY 830 uLysLeuCySGlnGlnLeuPheGlnLeuLysGlnSerLeuSerAspAlaGlnAlaValTh 850
 Db 2720 AAACGACTGTGAAGAGTTGGTGCAAATCAAGAGAAATAGAA-----GAAATCTCAT 2773
 QY 850 rArgAspAlaGlnLysGlnLysSerPheLeuArgSerGlnuAsnleuGlnleuLysGlnLys 870
 Db 2774 GAAAGCAGAACAGATGCATCAAAAGTTTGTGGCTGAAACAGATCAGCGCATTAAGTT 2833
 QY 870 smetGlnAspThrSerAsnTrpTyrAsn-----G 880
 Db 2834 ACAGAGAGACACTCTCTGCTCCAGAAATGTTGCTGGAACCTTAAGTCCCTTGAGAA 2893
 QY 880 nLysGlnLysAlaAlaSerLeuPheGlnLysGlnleuGlnThrlYsSerAsnTrpLys 900
 Db 2894 CAAGAGAAAGAGCTGCACACTTTTAATGATAGATGAAGAACTGACGACGACAGATGCA 2953
 QY 900 sLysMetGln-----AlaAspLeuGlnLysGlnleuGln 911
 Db 2954 AGAATTAAAAAGACCAACCATCTTACTGGAAGCTCTCTTAAGAGCTCAACCTTTATTC 3013
 QY 912 -----SerAl 913
 Db 3014 CGAAACCTTAAGCTTGAGAGAAAGAAATGATTCATCATTTCTTAATATAAGGGA 3073
 QY 913 apheAsnGlnLysAsnTrpLeuAsnGlnLysLeu-----AlaGlyLys 927
 Db 3074 AATTGAAGACTGACCCAGAGAAATGGAGCTCTTAAGAAATTAATGATCTCTTAATCA 3133
 QY 927 sValProArgAspLeuLeuSerArgVal-----GlnleuG 939
 Db 3134 AGACAGATGAACTTAATCCAGAAAGTGCAGAGTTTGCACAACTATATATAGTGAAGGGA 3193
 QY 939 uLysLysValSerGlnPheSerLysGlnleuGln----- 950
 Db 3194 GAAAGAGATTTTCAGAGTTATCTGATCAGTACAGCAAGAAACTTATTTACTACAAAG 3253
 QY 951 -----LysAlaLeuGlnG 955
 Db 3254 ATGTGAAGAAACCGGAATGCATATGAGATCTTATGTCAAAATATACAAAGACAGCA 3313
 QY 955 uLysAsnAla-----LeuGlnuAsnGlnValThrlYsLeuSerGlnuTrpLysPh 971
 Db 3314 AAAGAAATTTCTAAATTTAGAAATGCTTGTCTTAATGATGCACTAGCTGTGTGAAAAATAG- 3371
 QY 971 eleuProAsnGlnValGlnLysLeuLysAsnGlnLysSerLysAlaSerGlnGlnLysleu 991
 Db 3372 -----AAAAATGAGTTGAGAACGCTAAAGAGAAAGATTTGCAAGAGAACACAGAAATCTT 3427
 QY 991 tLeu---LeuLysGlnLysGlnLysSerAlaSerLysLysSerLysGlnLysleu 1010
 Db 3428 AACAAATTTAGCATTTCTGTGAGAAAGAAATCAAGAAATCTGATCTAGAGTTGAGAGAGT 3487

QY 1010 eMetGln-----GlnGlnSerGln----- 1016
 Db 3488 GCAGAACCTGTGAGATCTGAGATGACAGATTAACCAAAATCTTAAGACGAGGCTGC 3547
 QY 1017 -----GlnlleuGlnleuThraspGlnValThrlYsThrlYsSerLysVal 1032
 Db 3548 TGGTTAAAGCAAGAAATCATGACTTTAAAGAGAACCAAAATATGCAAAAGAGAGT 3607
 QY 1032 1-----GlnGlnThrlYsGlnLysGlnLysMetLysAsp-- 1047
 Db 3608 TAATGACTTATTAACAAGAAATGAAACAGCTGATGAAGTAATGAAGAACTTAACATGAATG 3667
 QY 1048 ----AspLeuPheGlnLysTrpLysAsnLysSerGlnLysAlaGlnAspLeuLysArg 1066
 Db 3668 TCAAAATCTTAATCAAGAACCAATTAAGAACTCTGTGAAAGAAAGAGAGTGAAGAA 3727
 QY 1066 uMetGlnuAsnleuLysGlnLysMetGln----- 1075
 Db 3728 TCAATGTATTTTAAACCTCAGATGATCTTGAAGTTAAAGAAATTTCTAGATCTTA 3787
 QY 1076 -----SerValGlnValLyslleAlaAspThrlYsHlsGlnleuGln 1090
 Db 3788 TAATGCCGAGTTGGTGCATTAAGAGCTATGCTTAAGAAATTAAGAAATTAACCTTACAG 3847
 QY 1090 uThrlleuArgAspLysGlnLysGlnleuHlsGlnLysLysTrpPhePheGlnAlaMetG 1110
 Db 3848 AAGTGAGAGAGAGAGAGGAGTCCCTGCACATGA-----TTACA 3886
 QY 1110 nThrlleuPheProIleuThrlYsLeuSerAspSerLeuProSerLysLeuValGln 1130
 Db 3887 GACATTT-----AGAGAGATCTTAAACCAACATTTTGCAGAGAT 3928
 QY 1130 YAsnSerGlnAspProIleuGlnLysAsnAspTrp----- 1141
 Db 3929 GCAGTCAACAAAGAAATTAAGTGCCTTAAGAGCTGGAATTAAGTGCAGAAAGATATAT 3988
 QY 1142 -----HlsAsnleuLyslleAlaLeuAlaThrlYsGlnArgAsnHlsMetValCys 1158
 Db 3989 TTCAGGCCCTCATGATGATTTGACACAGTCAAAAGACATGCACACCTTCAGTCTCTCT 4048
 QY 1158 uGlnThrlYsGlnArgAsnSerLysGlnLysGlnVallleAspLeuAsnThrlYsGln 1178
 Db 4049 GCAAAACATCAATGAAGCAAGCTGAATGAG-----CTAGAGAAATATGGAAT 4096
 QY 1178 rLeuGlnAlaGlnSerlleGln-----LysSerAspLeuG 1190
 Db 4097 ACTGCAGCTGAAAGATATGAACTGTAAGTGAAGTGAATGATTCAGAGTCAAGATGAT 4156
 QY 1190 nLysProLysGlnAspLeuGlnLysGlnVallleAspLeuLeuGlnMetGlnleu 1210
 Db 4157 CACAGCAACTGGAAGAAATGCGAGAGAGGTAGGAACTTAATTAATGAATTAATAT 4216
 QY 1210 uLysGln----- 1212
 Db 4217 AAATGATGACAGTGTCTTTCATGTGTGATTAAGTGAAGACATTAACAGAGGTGAAT 4276
 QY 1212 ----- 1212
 Db 4277 TGTGAAACAAATGAAACAGACCCGTGTCTTGGCTTCATTTGACGAGAGTAATTC 4336
 QY 1213 -----HlsleuThrlYsSerGlnleuSerlleGlnLysGlnleuGlnuAsnleuG 1230
 Db 4337 CTAGAGACACTTGCACA-----TTGTCAAGCAAAAGAAATTAATGATGCACTTGCAG 4387
 QY 1230 uValThrlYsGlnleuGlnThrlYsGlnLysGlnLysMetLysAsnHlsThrlYsGlnArg 1250
 Db 4388 ATTCAGAGAGAAATCTTATCTTAACAAGTGAACACAAATTTTATCATGATGACGACTG 4447
 QY 1250 nglnleuGlnThrasnPhesGlnAspLeuLysAlaGlnHlsAspSerLysGlnAspLe 1270
 Db 4448 TCAGATGAGCTCTTAATATGTCAGAGCTGCAGACCTATGTTGATTCATTAAG----- 4499

QY 1270 usergluasnileglugluserilegluuhrglnaspgluuvalaialieglucl 1290
 Db 4500 -GCCGAAATTTGGTCTTGCA -ACGAAATCTGCAAACTTCAAGTGA 4546
 QY 1290 uleuargluginlysleuvalaspserephargluginleuaspyserva 1310
 Db 4547 CTTCGTGAGAGAGATGCACTG -GGCTTGAGAGAGGGCTGCTTCCTCTTC 4600
 QY 1310 lgllylserseproasnhsaspalaialaasnlglnlulvalserleuyl 1330
 Db 4601 ATCTCTTGTGTGCTTACAGCTCTAGCTTACG -AGTTGGGAGA 4645
 QY 1330 uvalasnerleuglnserglumetleuarglgluargaspcluleuglnhrsercy 1350
 Db 4646 C--TCCTCTTTACAGAGCTCTTTAGAA -CAGACAGAGA 4684
 QY 1350 slyalaleuvalsergluleuglu - 1358
 Db 4685 TATGCTCTTTGAGTAAATTTAGAAAGGGCTTTACAGAAACAGTCAGTATGA 4744
 QY 1359 -Leuuaralalaasvalyservaigluylgluasnleuglu 1374
 Db 4745 AGTATTTGACAGAGCTGACAGCTGATGCTTAAAGCCGAAATTTGCTT 4804
 QY 1374 ethrlsysleuasnlglyleuglulysglulileugllylserglu -G1 1391
 Db 4805 GTCAACGAACTGAGAACTTTCAGAGGTGACTTGTGAGAGAGATGCACTGGCTTGA 4864
 QY 1391 usergluvalleuysersermetleugluasnleuylspsasnlylserleuys 1410
 Db 4865 GGAGGGGCTGCTTCATCCCTGTCATCTGTGTGCTGACAGCTCTAGCTTACAG 4924
 QY 1411 -gluglnlaleugluylser - 1417
 Db 4925 TTTGGAGACTCTCTTTTACAGAGCTCTTTAGAACAGCAGAGATATGCTCTTT 4984
 QY 1418 -serlygluasnlglnpbeserleugluvalphe - 1429
 Db 4985 GAGTAATTTAGAAAGGGTCTTTCAGCAACAGTCAGTATGAAATTTTGCAG 5044
 QY 1430 -serglyserglnlyleuva 1436
 Db 5045 CAGTCTGACAGAGAAATCTGACAGAAACCCCTTGCCGACGAGAGAGGT 5104
 QY 1436 laspgluile -gluvalleuysalaglnleuylsalalaglnluar 1452
 Db 5105 TGAAGCTTGAGCTCCCTGTGAGGTGACGGAGCTCCGAGAGAGCTAGAGAGA 5164
 QY 1452 gleuglu -ileuysaspargaspyrrhnegluleuvalglnthrilaas 1468
 Db 5165 AATGGAAGTCAAGGATTAATAAATGAGAAATTCAGAGCTCGAGCACTTTAAG 5224
 QY 1468 ntrhranleuvalglu -glylsleugluuhrrproleuglnlaasphsleu - 1485
 Db 5225 TTCTGAAAGCAAGAGCTTGACTGTAGGAAGCAAGTATTTGTCAAAAATGAGACAG 5284
 QY 1486 -gluaspsereileaspargaspyrrhnegluleuylvalleugllylserle 1505
 Db 5285 GCAACAGAACTGACAAAGCTGCTGAGATGAGTCCAAAGTGGCGGCAAAAAGA 5344
 QY 1505 ugluargasnlglnlyrleuuglu -Argleuglnlglnlylserleuyl 1521
 Db 5345 ACAGAGCAAGCAACTGTCACTTGAAGTGAAGTGAAGCAAGCTCAGAGTGTGA 5404
 QY 1521 uleuserasnlyleuugluileuglnlysglmetgluuhrrservalleuylsps 1541
 Db 5405 CTTRAGCTTCGG -TCTTGTGCTGGCAT 5431
 QY 1541 paspleuglnlylserleuglnserleuasnlelleleuylsasnli 1561
 Db 5432 CGACACAGAGAGCTATTCAGAGC -CGAAATGAGAGCTG 5470
 QY 1561 easprthrleuylshisshaserasprthrlnalaglnleuglnly 1577

Db 5471 TGACATATCAAAAGACATATCTTCAAGAACTACAGAAAGAACCAAGACATGATGTCA 5530
 QY 1578 -thrnglnlglnleuglnleu -Alaly 1586
 Db 5531 TCAGATTTGTATTAAGATGCTGACAGAGCTTCAATCTGACATTTGAGAAATTACTGA 5590
 QY 1586 sasleuualarlealalaseraasencyprolethrnglnlylserleu - 1604
 Db 5591 GACGTGCTCAGTGAACCCACAGAGAGAGTGTGGGACAGACCTCCCAATACCAATTA 5650
 QY 1605 -serlaspocysvalhlsproleu - 1612
 Db 5651 TGAGCTCCAGGGAGATTAACCCAGAGGCTCTTCAGAAATCTCTGAATTTGAT 5710
 QY 1613 -glululyslleleu 1618
 Db 5711 TTCTGCTCTAATGCTTTGTGATCCATATGATTTCTCTGGGAATCAGAAATATCTATA 5770
 QY 1618 uleuuhrgluginleuhsnglnlysthrasnlglnlulysleuuhisglulysas 1638
 Db 5771 TCTTCAACTGCGGTAAAGAGACATCAATGAGAAATTTGATTTACTTCTGTGATGA 5830
 QY 1638 ngluleuglnlalnlglnvalgluleuylcysgluvalglulhsleuylserme 1658
 Db 5831 GGACGTCAGCA -AAAGTTGAAAGTTGCTTAATGAAT 5869
 QY 1658 tleugluserlyserserleuglnserleuglnhlselulysasprhrngluginl 1678
 Db 5870 GAAAGAAATTAAGCTCAAACTC -CATTTACAGAGGTGCA 5908
 QY 1678 nleuleuala -leuylsglnlglnmetgluvalthrhl 1691
 Db 5909 ACTAATGACCAAAATTAAGACATGCAATGCAATTTGAAAAAATAGTTGGGAACTTAAGA 5968
 QY 1691 nglulysgluleuglnlthrhlsglnhlsleuuhlrhalaalvalasphlsleuyl 1711
 Db 5969 AGAAACTCAGATTTAAGTGAATAATTTCTGATTCATCCAGAGACTCT 6028
 QY 1711 sgluasnilegluleu -glyleuasnphelysasngluvalagln -G1lysthr 1728
 Db 6029 CCAGAGGTGAACCTTCTGAAGGCTCAATTCATTTAGAAATGATGAGATTAATTC 6088
 QY 1728 rthrlysglugincysleuasnlyl - 1739
 Db 6089 ATCACTGGAAGATATGAGATATATGCGCAAGGATGACAGCTGAGAGAGAT 6148
 QY 1740 -gluleuglnlserlglnhlsargleuglnlysglu - 1751
 Db 6149 TCTTGATGTGAATAAGCTGAGTATGATGATGCAAGGAAACCTGATTTGAAAA 6208
 QY 1752 -ilegluglnleuylserleuylsaspysgluseraleuuglu 1768
 Db 6209 AGCCCTTACTGTGAGGCTGACTTAGAGTATTCAAACAGAGAACTGTTTGAANA 6268
 QY 1768 rleuylserleuglnlyvalle -asnleuasnlglnlmetglumetvalme 1787
 Db 6269 AGACATGAATAATGACAGAGAGTATTTGCTCTTGAAGAAAGTCTCAGGTCTAC 6328
 QY 1787 tleuglnmetglululeuylsasnserrglnarthrvallealagluarg - 1804
 Db 6329 AAGTGAGAGAAACCGCTGTGTGAGAAATTAATGATCTATCTCAAAAAAACCGCAGCT 6388
 QY 1805 -aspglnleuglnaspsleuargluservalglumetserilegluuh - 1821
 Db 6389 GGATCAGTGTCTGAAAAATGAAAGAGAAACAA -GACCTTGAGTCTATCAAAAG 6445
 QY 1822 -glnaspsleuarglvalalaglnlale 1832
 Db 6446 TGAGTGTCTCATTCATTCAGTGCAGAGGAGAGAGGTGAAGAAAGACGAACTCT 6505
 QY 1832 uclnglnllyaspslysalnglnluleuuhrrserglnhlservalleuuglnly 1852

Db 6506 TCAGACTTTGCTCTGATGTAGTGCCTTTAAAGACAAATCATCTCCAGAGAA 6565
 QY 1852 sileserleuLeuGluAsnGlnMetleuTyrsValAlaThrValGluGluThrLeu 1872
 Db 6566 GCTGCAGACTTTGGAAAAGGACTCACAGGCACTGCTTTGACAAATGTGAGTGGAAAA 6625
 QY 1872 rgluArgAspPheLeuAsnGlnSerLeuGlnHisLeuPheSerGluIleGluThrLeu 1892
 Db 6626 CCAAAATGCACAACTGAAATTAAGAGAAAGAAATTCCTGTTCAGAAATTCGAAACCTGCA 6685
 QY 1892 rleuSerleuTySGluTySGluPhe-----AlaLeuGluGlnAl 1905
 Db 6686 GCCCAGACTGAGTATCATGCTTAAAGAGCTGATGCTCCAAAGCCCTTGAGGCGCC 6745
 QY 1905 aGluTySAspTySAlaAspAlaAlaArgTySThrIleAspIleThrGluTySLeuSer 1925
 Db 6746 ACTGCTGCAGAAAGGTGAGTTCGATTCGAGCTGAGCTCAGACACAGAGAGAGTGCATCA 6805
 QY 1925 nile-----GluGluGlnLeu 1931
 Db 6806 GCTGAGAAAGAGCATCGAAGAACTGAGACTTCGATTCAGCCGATGAAAGAGAGAGCT 6865
 QY 1931 uGlnGlnAlaThrAsnLeuTySGluThrLeuTySGluArgGluSerLeuIleGlnCysLy 1951
 Db 6866 GCACATTCGCGAGAAACTGAAAGAACGCGAGCGGAGATGATTCACCT-----AA 6916
 QY 1951 aGluGlnLeuAlaLeuAsnThrGlnHisLeuArgGluThrLeuTySLeuAspLeuAl 1971
 Db 6917 GGATTAAGTTGGAACTTGAAAGGAAATTCGATGTCAGAGAAAGAGAGAGTACT 6976
 QY 1971 aLeuTySGluMetGluGlnGluArgAspGluAlaAlaAsnTySValIleAlaLeuThrGlu 1991
 Db 6977 GAT-----CTTGATGCCGAGAAATTCAGACA----- 7004
 QY 1991 uTySmetSerSerleuGluGlnGlnLeuAsnValIleThrThrLeuTySGluGluGlu 2011
 Db 7005 -GAAATGAGACTCTAAAAACAATAATAGAGATGAGCGAGAGCCCTAAAGTTTTCATA 7063
 QY 2011 u-----GlyTySGluGluThrPheTySGluGlnArgProSerLy 2024
 Db 7064 ATTACACCTGTACAGTTAAGTCTGAAAAAGAAATTCGACAAAAACAATACAGAAAA 7123
 QY 2024 sGlnGlnSerSerGlnMetGluGluLeuArgGluSerLeuTySGluAspLeuGlu 2044
 Db 7124 ACAAGTCACTGTGAGAACTAGACAAAGTACTCTCTCATTTAAAGCTGTAGACAA 7183
 QY 2044 nLeuGluGluAlaGluTySGluIleSerGluAlaThr----- 2056
 Db 7184 AAAGAGACAGACAGATACAGATCAAAAGAAATCTAAACTGCAGTGCAGATGCTTCA 7243
 QY 2057 -AsnGluIleTySAsnLeuThrAlaTySLeuSerLeu-----GluGluGluIle 2074
 Db 7244 GAATAGTTAAAGAGCTTAATAGAGCACTACAGCCTTGCTGTGTCAGCAAGAAATTAAT 7303
 QY 2074 uGln-----AsnAlaSerIleLeuAsnGluAlaValSerGluArgGluAsnLeuArgHis 2093
 Db 7304 GAAGGCCACAGAAAGAGCTAGACCAACCATAGAGAAAGAGATCAGCTAGAAATAG 7363
 QY 2093 rTySGlnGlnLeuValSerGluLeuGlu-----GlnLeuSerLeuThrIle 2108
 Db 7364 CATTAAAGGCTGAGAGCCCGCTAGAGCTGATGAAAGAAAGAGCTGCTGTCTTACA 7423
 QY 2108 uTySAspArgAspHisAlaPheAlaGlnSerTySArgGluTySAspGluAlaValAla 2128
 Db 7424 ACAACTGAG-----GAAAGTGACATCATGACAGATTTACTTAAGGCTAG 7468
 QY 2128 sIleAlaSerLeuAlaGluGluIleTySLeuThrTySGluMetAsp----- 2144
 Db 7469 AGTGAGAACTTGAAAGAGACTAGAGATACCAAGAAACAAGACATGACAGCTCT 7528
 QY 2145 -GluPheArgAspSerTyS-----GluSerLeuGlnGlnGlnSerSerHisLeu 2161
 Db 7539 TGAGCGACAGAAATTCAAAAGAGAGTAGAGACCTTAAAGCAAAAATAGAAAGGATGAC 7588

QY 2161 rGluGluLeuCysThrTySGluThrGluLeuGlnMetleuTySGlnGlnAspIle 2181
 Db 7589 CCAAACTGTAGAGCTGTGGAAATTAGATGTTTACTATTAAGTCAAGAAAGAAATCT 7648
 QY 2181 eAsnAsnTySLeuAlaGluTySValTySGluValAspGluLeuGlnHisLeuSer 2201
 Db 7649 GACAAATGATTCAAAAGAGCAAGAGCAAGATATCTGAATTA----- 7691
 QY 2201 rleuTySGluGlnLeuAspGlnIleGlnMetGluLeuArgAsnGluTySLeuArgAsnTy 2221
 Db 7692 -----GAAATTAATAATTCATCATTTGAAAT-- 7718
 QY 2221 rGluLeuCysGluTySmetAspIleMetGluTySGluIleSerValLeuArgLeuMetGlu 2241
 Db 7719 -----ATTTTCAGAAAAAAGA 7735
 QY 2241 nAsnGluProGlnGlnGluTySAspValAlaGluArgMetAspIleLeuGluSerArg 2261
 Db 7736 GCAGAGAAAGTACAGATGAAAGAAATCAACACACTGCCATGAGATGCTTCAACA-- 7793
 QY 2261 gAsnGlnIleGlnGluMetGluTySLeuSerAlaValTySGluGlnHis 2281
 Db 7794 -----CAATTAAGACTCAATGAGAGAGTGCAGCCCTGCATTAATGACCA-- 7841
 QY 2281 rleuLeuSerSerLeuSerSerGluLeuGlnTySGluThrGluAlaHisLeuHisCysMe 2301
 Db 7842 -----GAAAGCTGTAG-- 7853
 QY 2301 rleuAsnIleTySGluSerLeuSerSerThrLeuSerArgSerPheGluSerLeuGlnThr 2321
 Db 7854 -----GCCAAAGACAGATATCTTAGTACAAAGTACAGAGTCTTGAACCT 7897
 QY 2321 rGluHisValTySLeuAsnThrGlnLeuGlnThrLeuLeuAsnLeuTySLeuValValTy 2341
 Db 7898 TGAGAAAGCTCACTGTCTCAAGCGCTTGAAGGCCAAAAAATTAATTAATGTTTTCGA 7957
 QY 2341 rArgThrAlaAlaValTySGluAspHisSerLeuIleTySAspTySGluTySAspLeuAl 2361
 Db 7958 ATCTTCAGG-----AATGCCATTCAGAAAGTGAAGATGCCAGCA 8002
 QY 2361 aAlaGluGlnTySArgHisAspGlu-----LeuArgLeuGlnLeuGlnCysLeuGlu 2378
 Db 8003 GAAATCGAGAAAGAAAGAAAGAAATCAGTACAGTGAAGAAATCAAAATTCAGAGCCAA 8062
 QY 2378 uGlnHisGluArgTySArgSerAspSerAlaSerGluGluLeuTySLeuPhe----- 2394
 Db 8063 GCAGCTTGTCTTAACTGTCCCGAGGTGAGAGAGAGACCAACTTGGAGAGAGCAAAA 8122
 QY 2395 -CysGluIleGluPheLeuAsnGluLeuLeuPheTySValAsnIleIleGlnSerVal 2414
 Db 8123 CTTAGAACTGAGAAATCTACAGCTGGAATTCGAGCGCAAGAAATTCAGATCTCAATTCGA 8182
 QY 2414 lGlnAspAspPheSerGluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlu 2434
 Db 8183 AAATGCCCTTTGACAGACACATTAGAGCTGCGAGAGTCTTACAGAAATCTAGAGAA 8242
 QY 2434 uGluLeuGlu-----HisTySLeuGluPheMetGlnThrPheGluGluPheGly----- 2450
 Db 8243 TGAGCTTGAATTCACAAAATGAGCAAAATGTCCTTTGTAAGAAAGTAAACAAATAGAC 8302
 QY 2451 -----AspLeuHisValAspAlaTySLeu-----Le 2459
 Db 8303 TGCAGAAAGAAACTGAGCTCAGAGGAAATTCATGAGATGCGACAGAAAAACAGACAGCT 8362
 QY 2459 uSerGluGluMetGlnGlnGluAsnArgGlnIleAlaSerThrIleGlnLeuLeuThrTy 2479
 Db 8363 GCAGAAAGAACTCAGTGCAGAGAAATTAAGCTGAGAGATTCAGTCTTACTGTTGGA 8422
 QY 2479 sArgLeuTySAlaValValGlnSerTySLeuGlnArgGluIleThrValTySLeuAsnGlu 2499
 Db 8423 AGAAATTAAGAGAC-----AGCAAAAGATCA-- 8447

OY	* 353	ThPrOHISVaLSnGLUValLeuAspSerGluAlaLeuLeuLysArgTyrArgysGlu	352
Db	112	ACAAGAGCTCTTCGAAAAATTTCAAGACCTTGAAGAGACACTTCACAAACTGTGAAGAGGA	171
OY	353	ILeUuAspLeuLysLysGlnLeuGluValnLeuGluSerSerSerGluThrLysAlaGln	372
Db	172	AAGCAGCAAAAGCGACTTTCAGCTTGCACAGCTCTCGAGGCTGCGCTCAGAGCAACACAG	231
OY	373	AlaMetAlaLysGluGlu-----HisThrGlnLeuLeuAla	384
Db	232	AAGGTGCAAAATGCAAAAAACCGAGGTACAAACCTGAAAAAGGAGCATCAAGATTGATG	291
OY	385	GluLeuLysGlnLeuHisLysGluValnArgGluAspArgIleThrHisLeuThrAsnIleVal	404
Db	292	GAAATATGTGAAAGCTCTGGAGAAACTAACAGACAAATTTCTCAT-----GAACCTTCA	345
OY	405	ValAlaSerSerGlnGluSerGlnAsp-----GlnArg	416
Db	346	GTCAAAGAGCTCACAAAGATTTCCAGGAAGGACAACTGAATTACAGGCCAAAAACAATA	405
OY	417	ValLysArgLysArgValIleThrPalaProGlyLysIleGlnAsnSerLeuHisAla	436
Db	406	GAAAAACTGGAACGAGACTTAAAGGTGAATTCTGACCTTGAAGAAGCCACAAACCT	465
OY	437	SerGlyValSerAspPheAspMet-----Leu	445
Db	466	GCGCAGTCTGCGAGATGTCTCTCGAATTCATGCATATACACCAACCAAAAAATTTTCAACT	525
OY	446	SerArgLeuProGlyAsnPheSerLysLysAlaLysPheSerAspMetProSer-----	463
Db	526	CCACTAACCCCAACTCAATTTATAGTGTTCACAGTATGAAGATCTAAAGAAAAATAT	585
OY	464	PheProGluIleAspAspSer-----ValCysThrGluPheSer	476
Db	586	AATTAAGAGGTGAAGAACCAAAAAAGTTTAGAGGCGAGAGTTAAAGCCTTCGACGCTTAA	645
OY	477	AspPheAspPalaLeuSerMetLeuAspSerAsnGlyIleAspAlaGluTriPasnLeu	496
Db	646	AAACCAACCAAGCACTTCCACAAAGCCACCATGAAATCACCGGACATTGGCCGCGATCAG	705
OY	497	AlaSerLysValThr-----HisArgGluLysThrSerLeuHis-----	509
Db	706	GCTTCATCATCTGTGTTCTTCATGCGACAGCAAGAAAGACCCCAAGTCACTTTTCATCTAAT	765
OY	509	-----	509
Db	766	TCTCAAAAGACTCCAAATTAGAGAGATTTTCTCGCATCTTACTTTTCTGGGACCAAG	825
OY	510	-----GlnSerMetIleAspPheGlyGlnIleSerAspSerValGlnPheHis	525
Db	826	GTGACTCCAAAGTGCATCACTTTTGCAAAATAGGAAAGAGATGCAATTAAGCAGTTTCTTT	885
OY	526	AspSerSerLysGluAsnGlnLeuGlnIleThrSerLeuGlnGlnLeuGlnSerLys	545
Db	886	GACATTTCTGACAGCTCTCATCTTTTG-----GATCAATTAATAA	924
OY	546	CysArgLysAlaSerPheGluLysGluIleThrSerLeuGlnGlnLeuGlnSerLys	565
Db	925	GCGCAGCATCAAGAGTAAGAAACAGATTATAGATTGAACTAGGACCTGCAAGACAT	984
OY	566	GluGlnGluLysLysGluLeuValGlnSerPheGluLeuLysIleAlaGlnLeuGlnGlu	585
Db	985	GAAAAACAAATGAAAGGCCCAAGTGAATTAAGTTTCA-----GAACCTCAACTC	1032
OY	586	GlnLeuSerValLysAlaLys-----AsnLeuGluMetValThrAsnSer---	600
Db	1033	CAACTGAG-----AAAGCAAAAGTGAATTAATTTGAAAAAGAAAGTTTGGACAATAATGT	1089
OY	601	ArgLinhIISerIleAsnAlaGluValGlnThrAspValGluLysGluValValArgLys	620
Db	1090	AGGATCACTAGTGAGAACACACAGCACAATCTAGC-----	1125

QY	621	GlutetservalLeuGlyAspSerGlyTyrAsnAlaSerAsnserAspleuGlnAspSer	640
Db	1126	-----CAGCGCTCAACCAAGTATCTGATTGGACAAACAAATGAAAAATTG	11737
QY	641	SerValAspGlyLysArgLeuSerSerSerHisAspGluCysIleGluHisArgLysMet	660
Db	1174	ACGGAAGAT-----TTGAGTTGTCACCGACGCAAAATGCAAGAAAGTCCAGATGTCTT	12244
QY	661	LeuGluGlnLysIleValAspleuGlu--GluPheIleGluAsnLeuAsnLysLysSer	679
Db	1225	CTGGAAACAGAAATTATAGGAAAAAGAAAGAGATTTCAGAGAGAGCTCCCGCT-----	12787
QY	680	GluAsnAspLysGlnLysSerSerGluGlnAspPheMetGluSerIleGluLeuCysGlu	699
Db	1279	-----CAACGACGCTTCTTCCAAACACAGCTGACAGAGAGTGCATCCAGATG-----	13223
QY	700	AlaIleMetAlaGluLysAlaAsnAlaLeuGluGluLeuAlaLeuMetArgAsnPhe	719
Db	1324	-----MAGCCAGACTACCCAGAGATTACAGCAAGCCAAAG---AATTG	13656
QY	720	AspAsnIleIle-----LeuGluAsn	726
Db	1366	CACACGCTGCTGCAGAGCTGAACCTGGATAACTCAGATCAGTAAAGCAACAGCTAGAAAAC	14253
QY	727	-----GluThrLeuLysArgGluIleAlaAspleuGluArgSer-----	739
Db	1426	AATTGGAAAGATTATTAAGCAAAATGTTGCGAGCTGCAACAGCGCTTCCAGCGAGTGC	14858
QY	740	LeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGluLysGluThrGln-----	757
Db	1466	ATCAAGAGCATGACCTGAGAGAGAGAGATGAGAGAAATGACAGAAACAAACCTCCTT	15455
QY	758	LysGluHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn	777
Db	1546	AAGAGTCACTCTGCAGAAAAGCCAGAGAGATGCTGCCACCTGGAGGCAACACTCAAGAAC	16050
QY	778	AlaGluMet-----TyrAsnGlnAsnLeuGluAspleuGluThrLysThrLys	794
Db	1606	ATCAACAGCTGTTTAATCAGAGCCCAAAATTTGCAAGAAATGAAACGCAAGAT---	16623
QY	795	LeuLeuLysGluGlnGluIleGlnLeuAlaGluLeuArgLysArgAlaAsnLeuGln	814
Db	1663	-----ACCTCTCGAGAAACCACTGTTAAGAGATCTTCAAGAAAAATTAATATGACAAAG	17166
QY	815	LysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysLeu-----	832
Db	1717	AACTCCTTGACTTTAGAAAAACGTAAGCTGCTGCTGCTGATCTGAAAAAGCAGCAAT	17766
QY	833	CysGluGlu-----IlePheGlnLeuLysGlnSerLeu	843
Db	1777	TGTTCTCAAGACCTTTTGAGAAAAAGAACATCATCATTTGAACAACCTTAATGATMAAGTTA	18366
QY	844	SerAspAlaGluAlaValThrArgAspAlaGlnLysGluGlySerPheLeuArgSerGlu	863
Db	1837	AGCAAGACAGAGAAAGAGCTCAAA--GCCCTGCTAGTGCCTTAGAGATTAAAAAGAAA	18933
QY	864	AsnLeuGlnLeuLysGlnLysMetGluAspThrSerAsnTrpTyrAsnGlnLysGlnLys	883
Db	1894	GAATATGATTAATGAAGAAGAAACAACTGTTCTTCTTGCGAAAAAGTCAAAACGAAAAA	19533
QY	884	AlaAlaSerLeuPheGluLysGlnLeuGluThrGluLysSerAsn-----	898
Db	1954	CTTTTAAC-----CAGATGGAATCAGAAAAGAAAACCTGCAGAGTMAATT	20011
QY	899	-----TyrLysLysMet	902
Db	2002	AATCACTTGGAACTGTCTGAAAGACACAGCAATAAAAAGTCATGAATACACAGAGGA	20611
QY	903	GluAlaAspleuGlnLysGlnLeuGlnSerAlaPheAsnGluIleAsnTyrLeuAsnGly	922
Db	2062	GTAAGAACGCTGCGAGATGACACAGAAACCTTAAGTCTGAGATCAGAAACCTTCAACAC	21211
QY	923	LeuLeuAlaGlyLys-----ValProArgAspLeuLeuSerArgValGluLeuGlu	939

Db	2122	GTGTTAACAAGTAAAGTCAGTGGAGGTGAAGAGCCAGAAACAGTACTTATATGAGGTACAG	2181
Qy	940	LysLysValSerLysPheSerLysGlnLeuLysAlaLeuGlnGluLysAsnAlaLeu	959
Db	2182	CAGAAA---GCTGAGTTCGA---GATCGAAAACATCGAAG	2217
Qy	960	GlnuGlnValAlaThrCysLeuSerGluYrLysPheLeuProAsnGlnValGluCysLeu	979
Db	2218	GAATATGAAATAATGTGTTTGAAGACTTCAG---CTTACTGGCGAAGTTGAAGATCTTA	2274
Qy	980	LysAsnGlnIleSerLysAlaSerGlnIleLysT---	991
Db	2275	GAACACAAGCTTCAGTTACTGTCGAATGAATAAATGAACAAGACCGGTGTACCAAGAC	2334
Qy	992	-----LeuLeuLysGlnGluGly	997
Db	2335	TTTGCATGCCGAATATGAGAGCTCAGAGGATCTGTCTAAATCCAAAGATGCTTCCTGGTG	2394
Qy	998	-----GluHis	999
Db	2395	ACAATATGAAGTATCATCAGAGAGCTTTTGGCTTTTGATCAGACAGCTGCCATGATCAT	2454
Qy	1000	Ser---AlaSerIleIleSerLysGlnIleIle	1010
Db	2455	TCCTTTGCATAATTAATTTGGAGAACAAGACATGCTTCAGAGAGAGTGAATGTCT	2514
Qy	1010	-----	1010
Db	2515	TTAGAGCAGACCAAACTCCGAAAATTTGGCATTCTCAAAAATAGAGTTGATTCACT	2574
Qy	1011	-----MetGlnGlnGlnSerGlnGln	1017
Db	2575	GAATTTTCATTAGAGTCTCAAAAACAGATGAATCAGACCTCAGCTCCCAAAACAGCTGGAAG	2634
Qy	1018	IleLeuGlnLeuThrAspGlnValAlaThrHisThrGlnSerLysValGlnGlnGluGln	1037
Db	2635	TGTGTGCAAATCAAGAGGAATATGAAGAAAATTCATGAAGCAAGACAGATGATCACTCA	2694
Qy	1038	GlnYrLeu-----GluMetLysLysMetHisAspAsp	1048
Db	2695	AGTTTGTGGCTGAACAAGTCAGCGCATTTAGTAAGTTTACAGGAAGACATCTTGCCTCAC	2754
Qy	1048	-----	1048
Db	2755	CAGAGTGTGTGCTGCTGAACCTTAAGTGCCTTGAAACACAGAAAAGAGCTGCACCT	2814
Qy	1049	LeuPheGlnLysTyrIleArgAsnLysSerGln	1059
Db	2815	TTAAATGATTAAGGTATGAAGTCAAGCAGCAGAGATTTCAAGAAATTAAGAGACACCAT	2874
Qy	1060	---AlaGluAspLeuLeuArgGlnMetGlnAsnLeuLysGlyThrMetGlnSerValGln	1078
Db	2875	CTACTTGAAGACTCTTAAGAGAGCTCAACCTTTATCCGAAAACCTTA---AGCTTGAG	2931
Qy	1079	ValLys-----IleAlaAspThrLysHisGlnLeuGlnGlnAlaThrIleArg	1093
Db	2932	AGAAAGAAATGATAGTTCATCATTTCTCTAAATAAAGGAATAATGAAGAGCTGACCCAA	2991
Qy	1094	Asp-----LysGln-----GlnLeuLeuHisGlnLysLysTyrPhePhe	1106
Db	2992	GAGAAATGGACTCTTAAGAGAAATTAATGATCCTTAAATCAAGAGAAAGATGAATTAATC	3051
Qy	1107	GlnAlaMetGlnThrIlePhe-----ProIleThrProIleu	1118
Db	3052	CAGAAAAGTGAAGCTTTGCAAACTATATAGATGAAGGAGAAAGCATTTTCACAGATTA	3111
Qy	1119	SerAspSerLeuProProSerLysLeuVal-----GlnGlyAsn	1131
Db	3112	TCTGATCAGTACACAGAAAACCTATTTTACTACAAAGATGTGACAAAGACGGAAAT	3171
Qy	1132	SerGlnAspProIle-----	1136

Db	3172	GCATATGAGGATCTTGTAGTCAAAAATATCAAAAGCAGCAGCAAGAAAAAGATTCTAAATTAGAA	3231
Qy	1136	-----	1136
Db	3232	TGCTTGCTAAATGAAATGACACTAGCTCTTGTGAAATATGAAAAATGAGTTGGAAACGTA	3291
Qy	1137	-----GluTLeAsnAspTyrHisAsnLeuIleValLeuAlaThrGlu	1150
Db	3292	AAGGAGCAATTTGCGAAGAGCAACCAAGATTTCTTAACAAATTAAGATTGGTGAAGAA	3351
Qy	1151	ArgAsn--AsnIleMetValCysLeuGluThr-----	1160
Db	3352	AGAAATCGAATCTGATGCTAGAGTTGGAGACGTGAGCAAGCTCTGAGATCTGAGATG	3411
Qy	1160	-----	1160
Db	3412	ACAGATATCCAAACAATTTTAAAGCCAGAGCTGGTGGTTTAAAGCAAAATCATGAGCT	3471
Qy	1161	-----GluArgAsnSerLeuLysGluValaIleAspLeu-----AsnThr	1174
Db	3472	TTAAAGGAAAGAAACAACAATGCAAAAGAAAGAAAGAAATTAAGCTTAATTAACAAGATGA	3531
Qy	1175	GlnLeu-----GlnSerLeuGlnAlaGlnSerIle	1184
Db	3532	CAGCTGATGAAGGTAATGAAGACTAAACATGAATGTCAAAATCTTAAGATCGAAACCAAT	3591
Qy	1185	GluLysSerAspLeuGln-----LysProLysGln	1194
Db	3592	AGGACCTCTGTGAAAGAAAGAGAGAGTGAAGAAATCAATGTAATTTAAACCTCAAGATG	3651
Qy	1195	AspLeuGluGluGluGluValLysLeu-----LeuLeuGlnMetGln	1208
Db	3652	GATCTTGAAAGTAAAGAAATTTCTCTAGATAGTTATATGCGACCTGGTGCATTTGAA	3711
Qy	1209	LeuLeuLysGlnLysLeuThrAspSerGlnLeuSerIleGluLysLeuGlnLeuLysAsn	1228
Db	3712	GCTATG-----CTAAGAAATTAAGAAATTTAAACTTCAGAAAGTGAAGAGAGAG	3762
Qy	1229	LeuGluValThrGluLysLeuGlnThrLeuGlnGlnMetLysAsnIleThrIleGlu	1248
Db	3763	-----GAGTGCCTGCACATGAATTTACACAAATT-----	3792
Qy	1249	ArgAsnGluLeuGlnThr--AsnIleGluAspLeuLysAlaGlnHisAspSer--Leu	1266
Db	3793	AGAGGAAATCTTGAACCCAGCAATTTGCAAGACATGCAAGTCAAGAAATTAAGTGGCTT	3852
Qy	1267	Lys-----GlnAspSerGlu	1272
Db	3853	AAAGACTGTGAATATGATGCGGAAAGAAAGATATTTCAGGGCTCATGAGTGTCAACA	3912
Qy	1273	AsnIleGlnInserIleGluThrGlnAspGluLeuArgAlaAlaGlnGluLeuArg	1292
Db	3913	AGTCAAAAGCAAAATGACACACCTTCAGTGGCTCTCAAAACAACATGAACAGCTGAAT	3972
Qy	1293	GluGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIle	1312
Db	3973	GACCTGAGAAATATGTGAATACTGACAGGCTGAAGAAAGTATGAATCTGTAACTGAGCTG	4032
Qy	1313	SerSerProAsnHisAspAlaValAlaAlaGlnIleLysValSer-----	1327
Db	4033	AATGATTCACAGTCCGAATGTATATGACACAGCAACTAGGAAATGGCAGAAAGAGGTAGGAAA	4092
Qy	1328	---LeuGluGluValAsnSerLeu-----GlnSerGlnMetLeuArgLysGluArgAsp	1344
Db	4093	CTCTTAATGAAGTTAAATATGTAAATGATGACAGTGGTCTTCTCCATGTGTGATTAAGT	4152
Qy	1345	Glu-----	1345
Db	4153	GAAAGACATACAGAGAGTGAATTTGGTGAACAAACCAATGAACAGCACCTGTGTCTTG	4212
Qy	1345	-----	1345
Db	4213	GCCTCATTTGGACAGAGATTAATCTCTAAGAGAGACTTGACATTTGCAGACAAAGAACTTCA	4272

QY 1346 -----LeuGlnThrSerCysLysAlaLeu 1353
DB 4273 ATGCACCTTGGCGAATTGCAGAGAAATTTCTTATTACAAAGGAGACACAAATTTTAA 4332
QY 1354 -----ValSerGluLeuGluLeuLeuArgAlaHisValLysSer 1366
DB 4333 CATGATCAGCACTGTAGATGAGCTTAATGCTCAGAGACCTGACACCTATGTTGACTCA 4392
QY 1367 ValGluGluGluAsnLeuGlnIleThrLysLysLeuAsnGlyLeuGlnLysGluIleLeu 1386
DB 4393 TTTAAAGCCGCAAAATTTGGTCTGTTCACACGAACTTGAGAAATTTCAAGGTGACTGTGTG 4452
QY 1387 GlyLysSerGlu-----GluSerGluValLeuLysSerMetLeuGluAsnLeuLys 1403
DB 4453 AAGAGATGCAAGCTGGGCTTGAGAGAGGGGCTGCTCCATCCCTGTCATCTCTTGTGTG 4512
QY 1404 GluAspAsnAsnLysLeuLysGluGlnAlaGlu----- 1414
DB 4513 CCTGACAGCTCTAGCTTAGCAGTTTGGAGACCTCCCTTTACAGAGCTCTTTAGAA 4572
QY 1415 -----GluTyrSerSerLysGluAsnGlnPhe 1423
DB 4573 CAGACAGAGATATCTCTTTTGTAGTAATTGAAGGGGCTGTTTCAGAAACCAAGTC 4632
QY 1424 SerLeuGluGluValPhe----- 1429
DB 4633 AGTGTAGATGAATTTTTCAGCAGCTGTGACAGAGGAACTTCACAGAAAGAAC 4692
QY 1430 ---SerGlySerGlnLysLeuValAspGluLe-----GluValLeuLysAla 1444
DB 4693 CCTTCGGCCCCAGAGGGGTGTGAAGAGCTTGTCCCTGTGAGCTGTACCGCAG 4752
QY 1445 GlnLeuLysAlaAlaGluGluArgLeuGlu-----IleLysAspArgAspTyr 1460
DB 4753 TCCTTCGAGAGCTAGAGAGAGAAATGGAAGTCAAGGATTAAGAAATTAAGAAATTT 4812
QY 1461 PheGluLeuValGlnThrAlaAsnThrAsnLeuValGlu---GlyLysLeuGlnThrPro 1479
DB 4813 CAAAGAGCTCGACAGTATTATTAAGTTCGAAAGCAAGAGCTTGAAGTTCAGCAAGAG 4872
QY 1480 LeuGlnAlaAspHisGlu-----GluAspSerIleAspArgArgSerGluGluMetGlu 1497
DB 4873 TATTTTCAGAAATGGAACAGTGGCACAGAACTGACAAACCTGAGCTGTGAGATGGAG 4932
QY 1498 IleLysValLeuGlnLysLysLeuGluArgAsnGlnIleThrLeuGlu----- 1513
DB 4933 TCCAAAGTTGGCGCAAGAAAGAAACGAGCAAACTGTCACTGAGCTGGAAGTACCA 4992
QY 1514 ArgLeuGlnGluGlnLysLeuGluLeuSerAsnLysLeuGlnIleLeuGlnLysGluMet 1533
DB 4993 CGACTCCAGCTACAAAGGTCTGACTTAAGTCTCGG----- 5028
QY 1534 GluThrSerValLeuLysAspAspLeuGlnGlnLysLeuGluSerLeuLeuSerGlu 1553
DB 5029 -----TCTTTCCTGGCATGCACACAGAAAGATGCTATTCAAGGC----- 5067
QY 1554 AsnIleIleLeuLysGluAsnIleAspThrThrLeuLysHisSerAspThrGlnAla 1573
DB 5068 -----CGAAATGAGAGCTGTGACATATCAAAAGAACATACTTCCAGAACTACAGAA 5118
QY 1574 GlnLeuGlnLys-----ThrGlnGlnGlnLeuGln 1583
DB 5119 AGAAACACCAAGCATGATGTTTCATCAGATTGTGATAAAGTGTCTCAGCAGAGCTCAT 5178
QY 1584 Leu-----AlaLysAsnLeuAlaIleAlaAlaSerAspAsnCysProIle 1598
DB 5179 CTACACATTTGAGAAATTAAGTGAAGCTGTGCTATTGAAACCCACAGAGAGAGTCTGGG 5238
QY 1599 ThrGlnGlnLysGluThr-----SerAla 1606
DB 5239 GAACAGTCCCAAGATACCAATTATGAGCTTCACAGGGAAGATTAACCCAGGCTCTTCA 5298

QY 1607 AspCysValHisProLeu----- 1612
DB 5299 GAATGCATTTTCGATTTGTCATTTTCTGCTGCTCATGCTTGTGCTATGATTTCTCTG 5358
QY 1613 -----GluGluLysIleLeuLeuLeuThrGluGluLeuHisGlnLysThrAsnGluGln 1630
DB 5359 GGGAAATCAGGAAGATATCCATATCTTCMACTGCGGGTAAAGAGACATCAATGAGAT 5418
QY 1631 GlyLysLysGlnGluLysAsnGluLeuGlnAlaGlnValGluLeuLysGlyGlu 1650
DB 5419 TTGACATTAATCTTCATGTATAGAGAGCCGTAGCAGA-----AAA 5457
QY 1651 ValGlnHisLeuMetLysSerMetIleGluSerLysSerSerLeuGluLeuGlnHis 1670
DB 5458 GTTGAAGTTTCTGTAATGAATGAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 5502
QY 1671 GluLysHisAspThrGluGlnGlnIleLeuAla-----LeuLys 1683
DB 5503 -----CATTTACAGAGAGTACAACTAAATGACAAATTTGAAGCATGATAGAAATGGAA 5556
QY 1684 GlnGlnMetGlnValValThrGlnGluLysGluGlnGlnIleThrHisGlnHisLeu 1703
DB 5557 AAAATAGTTGGGAACTTGAAGAAAGAACTCAGATTGAATGAATTAAGATTAATTTT 5616
QY 1704 ThrAlaGluValAspHisLeuLysGluAsnIleGluLe-----GlyLeuAsnPheLys 1721
DB 5617 TCTTGTGATCAACAGAGATTAATCTCAGAGAGTAAAGAACTTCTGAAGGCTCATTTCTGAT 5676
QY 1722 AsnGluAlaGln---GlnLysThrThrLysGluGlnCysLeuLeuAsnGluAsnLys--- 1739
DB 5677 TTGAAATGACATGACAGATTAATCATCAGCTGAAAGATTAATGAGATTAATGAGCTG 5736
QY 1740 -----GluLeuGlnGlnSerGlnHisArgLeuGlnLys 1750
DB 5737 AATGCACCTGGAAGAGATTTTCTGATGTGGAAATGAGCTGAGTACGATTCGATCG 5796
QY 1751 Glu-----IleGluGluLeuMetLysSerLeuLys 1760
DB 5797 GAGAAAGCTGACATGAGACATGAAGCCCTACCTCGAGGGGTGAGATTAAGATTAATTCGA 5856
QY 1761 AspLysGluSerAlaLeuGlnThrLeuLysGluSerGluGlnLysValIle---AsnLeu 1779
DB 5857 ACAGAGAACTATGTTTGAAGAAAGACATGAATTAAGACAGAAAGTATTTCTGCTCT 5916
QY 1780 AsnGlnGluMetGluMetValMetLeuGlnGluGlnGluGlnLysAsnSerGlnArgThr 1799
DB 5917 GAGAGAGACTCTCAGTGTGTGTCAGAGTGAAGAAACAGCTTCGTGGAAATTAAGTACT 5976
QY 1800 ValIleAlaGluArg-----AspGlnLeuGlnAspAspLeuArgGluSerValGlu 1816
DB 5977 ATGTCAAAATAACCAAGCAGCTGATCAGTTGTCTGAAAATAAGAGAGAAACACACA 6036
QY 1817 MetSerIleGluThr-----GlnAspAsp 1824
DB 6037 ---GAGCTTGAAGTCTCATCAAAAGTGAAGTGTCTCATTCATTCAGGTGCGCAGAGCAG 6093
QY 1825 LeuArgLysAlaGlnAlaLeuGlnGlnGlnLysAspLysValGlnGluLeuThrSer 1844
DB 6094 GTGAAGGAAAGAGCGAAGCTCTTCAGACTTTTCTGCTGTGATGTGAGTGAAGTGTAAAA 6153
QY 1845 GlnIleSerValLeuGlnGlnLysIleSerLeuLeuGlnGlnMetLeuTyrAsnVal 1864
DB 6154 GACAAATCTATCTCCAGAGAAAGCTGAGAGTTTGAAGAAAGAGCTCAGAGAGCTGTCT 6213
QY 1865 AlaThrValLysGluThrLeuSerGluArgAspAspLeuAsnGlnSerLysGlnHisLeu 1884
DB 6214 TTGACAAATGTGAGCTGGAACCAATTCACAACTGAATTAAGAGAAAGAAATTCCT 6273
QY 1885 PheSerGluIleGluThrLeuSerLeuSerLeuSerLysGluLysGluPhe----- 1900
DB 6274 GTCAAGAGATGTGAAGGCTCAGAGCCAGACTGAGTGAATGAGATTAAGAAACCTGAT 6333
QY 1901 -----AlaLeuGlnGlnAlaGluLysAspLysAlaAspAlaIleArgLysThrIle 1917

Db 6334 GCTCCAGGCGCTGGAGCCGCACTGGTGAGAGAAAGTAGTCCGATTTGAGGCTGAGC 6393
 Qy 1918 AspIleThrGluTyrLysIleSerAsnIle----- 1926
 Db 6394 TCAGACAGAGAGGAGCATCAGCTGAGAGAGGATCGAGAAAGCTGAGAGTTCGCAAT 6453
 Qy 1927 -----GluGluGlnLeuLeuGlnGlnAlaThrAsnLeuTyrGlu 1943
 Db 6454 GAGGCCGATAGAAAGAGCAGCTGCACATCGCAGAGAACTGAAAGAGCGGAGG 6513
 Qy 1944 ArgGluSerLeuIleGlnCysGluGlnLeuAlaLeuAsnThrGluHisLeuArgGlu 1963
 Db 6514 AAGGATTCACCTT-----AAGGATAAGGTGAGAACTTGAAAGGAAATGCGAGATG 6564
 Qy 1964 ThrLeuTyrSerLysAspLeuAlaLeuGlyLysMetGluGlnGluArgAspGluAla 1983
 Db 6565 TCAGAGAGAAACCGAGAGTAGTGATTT-----CTTGATCCCGAATTCGAAAGCA--- 6615
 Qy 1984 AsnLysValIleAlaLeuThrGluLysMetSerSerLeuGluGlnIleAsnGluAsn 2003
 Db 6616 -----GAGGTAGAGACTTAAACACAAATAGAGAGATG 6651
 Qy 2004 ValThrThrLeuLysGluGly-----GlyGluTyrGluThr 2016
 Db 6652 GCCAGAGCGCTGAAAGTTTGAATGACCTTGTCAGCTTAAGCTCGTGAAGAAAT 6711
 Qy 2017 PheTyrLeuGlnArgProSerLysGlnGlnSerSerGlnMetGluLeuArgGlu 2036
 Db 6712 CTGACAAACAAATACAGAAACAGGTCAGTTGTCAGAAACAGCAAGTTCCTCT 6771
 Qy 2037 SerLeuTyrThrLysAspLeuGlnLeuGluLysGluLysGluIleSerGluAlaThr 2056
 Db 6772 TCATTAAAGTCTGTATGAGAAAGAGAGCAGACAGATACGATCAAGAAATCT 6831
 Qy 2057 -----AsnGluLysAsnLeuThrAlaLysIleSerSer 2068
 Db 6832 AAAATCGACGTGAGATGCTTCAGAAATCAGTTAAAGAGCTAAATGAGCGAGTACGCC 6891
 Qy 2069 Leu-----GluGluGlnIleLeuGln-----AsnAlaSerIleLeuAsnGluAlaValSer 2085
 Db 6892 TTGCTGTGTGACCAAGAAATATATAGAGCCACAGACAGATCTAGACCCACCAATAGAG 6951
 Qy 2086 GluArgLeuAsnLeuArgHisSerLysGlnGlnLeuValSerGluLeu----- 2102
 Db 6952 GAAGAGCATCAGCTGAAATATGATTCGAAAGCTGAGAGCCCGCTAGAGAGCTGATCAA 7011
 Qy 2103 -----GlnLeuSerLeuThrLeuLysSerArgHisAlaPheAlaGlnSerLysArg 2120
 Db 7012 AAGAGCGCTGCTGTCTTACAACTGAG-----GAAAGTGAGCAT 7056
 Qy 2121 GluLysAspGluAlaValAsnLysIleAlaSerLeuAlaGluIleLysIleLeuThr 2140
 Db 7057 CATGCAGATTACTTAAGGATAGAGTGGAGAACCTTGAAAGAGAGCTAGATAGCCAGG 7116
 Qy 2141 LysGluMetAsp-----GluPheArgAspSerLys-----GluSerLeu 2153
 Db 7117 ACAAAACCAAGAGCATGACGCTGTGAGCGACAGAAATTCGAAAGAGAGAGTGAACCTTA 7176
 Qy 2154 GlnGluGlnSerHisLeuSerGlnGluLeuLysSerThrTyrLysThrGlnLeuGlnMet 2173
 Db 7177 AAGCAAAATAGAAAGGATAGCCCAAGTCTGAGAGGCTGAGATTGATGATTGTTACT 7236
 Qy 2174 LeuLysGlnGlnLysGluAspIleAsnAsnLysLeuAlaGluLysValLysGluValAsp 2193
 Db 7237 ATAAGGTGAGAAAGAAATCTGACAAATGATATACAAAGAGAGCAAGCAGATATCT 7296
 Qy 2194 GluLeuLeuGlnHisLeuSerSerLeuLysGluGlnLeuAspGlnIleGlnMetGluLeu 2213
 Db 7297 GAATTA-----GAAATTA 7308
 Qy 2214 ArgAsnGluLysLeuArgAsnTyrGluLeuLysGluLysMetAspIleMetGluLysGlu 2233

Db 7309 ATAAATTCATCATTTGAAAT----- 7329
 Qy 2234 ILeSerValLeuArgLeuMetGlnAsnGluProGlnGlnGluAspValAlaGlu 2253
 Db 7330 -----ATTTCGCAAGAAAGAGAGAGAGAGATGACATGAAAGAAATCAAGCACT 7383
 Qy 2254 ArgMetAspIleLeuGlnSerArgAsnGlnGluIleGlnGluLeuMetGluLysIleSer 2273
 Db 7384 GCCATGAGATGCTTCAACA-----CAATTAAAGAGCTCAATGAGAGATGGCA 7434
 Qy 2274 AlaValTyrSerGluGlnHisThrLeuLeuSerSerLeuSerGlnLeuGlnLysGlu 2293
 Db 7435 GCCCTGCAATAGCCACA----- 7452
 Qy 2294 ThrGluAlaHisLysHisCysMetLeuAsnIleLysGluSerLeuSerThrLeuSer 2313
 Db 7453 -----GAAGCCTGTAAG-----GCCAAGAGCAGAAATCTTAT 7485
 Qy 2314 ArgSerPheGlySerLeuGlnThrGluHisValLysLeuAsnThrGlnLeuGlnThrLeu 2333
 Db 7486 AGTCAGTAGAGTGTCTGTAACCTTGAAAGGCTGACAGCTTCAAGGCTGATGAGGCC 7545
 Qy 2334 LeuAsnLysPheLysValValTyrArgThrAlaAlaValLysGluAspHisSerLeuIle 2353
 Db 7546 AAAAATATATATATTTTGTTCATCTCACTG-----ATAGCCTCAT 7590
 Qy 2354 LysAspTyrGluLysAspLeuAlaAlaGluGlnLysArgHisAspGlu-----Leu 2370
 Db 7591 CAAGAGTAGAATAGATGCAAGCAAGAACTGAGAGAAAGATGCAAAATCAGTAGACTG 7650
 Qy 2371 ArgLeuGlnLeuGlnCysLeuGlnGlnHisGlyArgLysTyrSerSerPheAlaSerGlu 2390
 Db 7651 AAAATCAATTCGAAACCAAGAGCAGCTTGCTCAACTGCTGACGAGTGAAGAGAG 7710
 Qy 2391 GluLeuLysPhe-----CysGluIleGlnPheLeuAsnGluLeuLeuPheLys 2406
 Db 7711 CACCACCTTGGAGAGAGCAAAACTAGAACTAGAAATCTGACAGTGAATGGAGCAG 7770
 Qy 2407 LysAlaAsnIleIleGlnSerValGlnAspAspPheSerGluValGlnValPheLeuAsn 2426
 Db 7771 AATGTCAGAGCTGACAAATCCAAATGCTCTTCAGAGACACATAGAAAGTGTGAG 7830
 Qy 2427 GlnValGlySerThrLeuGlnGlnGluLeuGlu-----HisLysLysGlyPheMetGlnTyr 2445
 Db 7831 AGTTCTTACAGAAATCTAGGAATGAGCTTGAATTCACAAATATGGACAAATGTCTCT 7890
 Qy 2446 LeuGlnGluPheGly-----AspLeuHisVal 2454
 Db 7891 GTTGAAAAAGTAAACAAATAGACTGCAGAAAGAACTGAGCTGCAGAGGAAATGCTAGAG 7950
 Qy 2455 AspAlaLysLys-----LeuSerGlnGlyMetGlnGlnLysAsnArgGluIleAla 2471
 Db 7951 ATGGCCAGCAAGAAACAGCAGAGCTGCAGAAAGAACTGAGGAGAAATATAGCTAGCT 8010
 Qy 2472 SerThrIleGlnLeuLeuThrLysArgLeuLysAlaValAlaGlnSerLysIleGlnArg 2491
 Db 8011 GGAGAGTTCAGATTACTGTTGGAAGAAATAAAGAGC-----AGCAAGATCA--- 8058
 Qy 2492 GluIleThrValTyrLeuAsnGlnPheGluAlaLysLeuGlnGlnLysGluGlnAsn 2511
 Db 8059 -----TTGAAGGAGCTGCACATCTGAAAT 8082
 Qy 2512 LysGluLeuMetArgArgMetGlnHisGlyProSerAlaSerValMetGluGlnGlu 2531
 Db 8083 AGTGAATTAAGAGAG-----AGCTAGATTGATGACCAAAAGAC 8121
 Qy 2532 AsnAlaArgLeuGlnGlyIleLeuLysThrValGlnAspLysSerLysLeuGlnSer 2551
 Db 8122 CAGGTGAGAAAGAGAGG-----AAGGTGAGAGAGAAATAGCTGATATACAGCTA 8172
 Qy 2552 ArgIleLysMetLeuGlnGlnGluLeuAsnLeuValLysAspAspAlaMetHisLysGly 2571
 Db 8173 CGG-----CTTCATGAGAGCT 8187


```

Oy 2572 GluIys-----ValAlaIleLeuGlnAspLysLeuLysArgAsnAlaGluIle 2589
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8188 GAAAGAAACACCGAGCTTGGTGGAC-----ACAAACAAACAGTAAAGTAGAA 8241
Oy 2590 LeuAsnAlaMetGlnValLysLeuThrLysGlnAspAsnLeuGlnAlaAlaMetLys 2609
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8242 ATCCAGACATACCGAGAAATAGTACTCTTAAGAAAGATGCTCAGTTCCAGAAAGCTG 8301
Oy 2610 GluIleGlnAsnLeuGlnLysMetValAlaLysGlnAlaValProLysGlnLys 2629
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8302 GAGATAGACCTTTAAAGTCTAGT-----AAAGAAAGAGCTC 8337
Oy 2630 Aspaen-----LeuLysThrLysValValLysIleLeuGlnIle 2642
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8338 AATATATTCATGAAAGCTACTACTACTGATTTGGAGAAATGAAACCAAGATGGAC 8397
Oy 2643 LysIleLysIle-----SerLysAlaThrAspGlnIle 2654
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8398 AATCTAAATATGTAATCAGTTGAAAGAAATGAAGCTGCCAGGGGAAATGAAG 8457
Oy 2655 AlaIleLysLysSerCys-----LeuGlnAspLysGlnLysIleLysArgLys 2672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8458 TTGTTGATCAATCTGTAACACGCTGGAAGAAAGAGATCTCCAGAAAGACTC 8517
Oy 2673 GluGlnLeuArgAlaGlnAlaAspAsnAspThrThrValCysValProLysAsp 2692
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8518 TCTCACTTCAAGCTGCACAGAGAGAGAGAAACA----- 8553
Oy 2693 GlnLysAlaSerThrPheProValThrCysGlnLysGlnLysIleValGlnSer 2712
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8554 -----GTTACGTTATGATGACC 8571
Oy 2713 AlaMetLeuValLeuGlnSerGlu-----LysAlaAlaLeuGln 2725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8572 AAGGTCGATGATTAACACAGTCAAGAACTGAAGAAACTCTGAAGAAAAAAC 8631
Oy 2726 ArgGlnLeuSerHisLysLysLysLysLysLysLysLysLysLysLysLys 2745
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8632 AAGGAGGAGATGATTAATCTGATTAAGTACTGTTCTCTG-----CTTATTAAGCAT 8682
Oy 2746 GluAspArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 2765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8683 GAAAGAGTAGAAGAAAGTAAAGAGATGTTAGACACAAAGTGGCCATCTGTTCA 8742
Oy 2766 His-----ArgGlnSerProHisLysThrGlnThrArgHisLysPro 2780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8743 CAATCTAAACAGATTCCTCCGAGGCTCTCT-----TTGCTAGGTCGA 8784
Oy 2781 ValThrProGlnArgSerGlnMetProSerLeu-----HisLeuGlnSerPro 2797
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8785 GTTGTCCAGAGACATCTCCATCTCTCTGTTACTGAAAGAGTTTATCATCTGGCCA 8844
Oy 2798 LysSerGlnSerSerThrLysArgValLys----- 2808
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8845 AATAAGGCTTCAGCAAGAGCAAGATCACTGATGATGAGAAATGTGGAGACCA 8904
Oy 2809 ---ProAsnArgSerGlnLysSer-----GlnLeuValMetSer 2821
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8905 ACACCTGCTACCCAGAGAGCTTTTCTAAAAAAGCAAGAACGATCATGAGTGTATT 8964
Oy 2822 ---ProGlnLys----- 2824
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8965 CACCTGAGAGAAAGACAGAGAGTACTGATTGAGCAGAGGACTTCCAGAAAGTTGTA 9024
Oy 2825 -----ThrGlnMetHisLysHisLys----- 2831
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9025 AAGAAAGGTTTCTGATCATCCGACAGAAAGACTACCATATATCTCTCGAAGAAC 9084
Oy 2832 -----LeuSerProSerLysValGlnLysLysLysLysLysLysLysLys 2847
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9085 ACCATGCACTCGAGACCAAGCCCGCTGCTGACAG-AAAGTATGCGTATCCCACT 9143

```

```

Oy 2848 -----AsnArgSe 2850
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9144 GAGTCTCGCAAGAAATCTTCAGAGTCTCCAAACCAAGCTGTGGCAGCATC 9203
Oy 2850 rGlnMetProThrGlnHisValLysSerProGlnLysThrGln-----LeuHisLysAs 2868
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9204 ACAAAGGTCAAAGTTGCTCAGCGAGGCCCACTAGATTGAGCAGACCATCTCCGAAACC 9263
Oy 2868 nLeuThrGlnSerThrLysPheAspAsnLeuSerProCysLysGlnLysValGln 2888
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9264 CACCAAGAAATCTCCAGTCAATATCTCTGAGAGAAAGTCCGACTGACAGCCGAG 9323
Oy 2888 nGlnAsnLeuAsnSerProLysGlnLysLeuPhe-----AspValLysSe 2903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9324 AGAGGCTCGAGGTCACAGCGAGGCCGACTTCTCCCAAGCCCAAGCTGAGTGGAGTC 9383
Oy 2903 rLysSerMetProLysCysProSerGln 2912
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9384 CAAGGCACTGAGAACTGTAAGTCCAG 9411

RESULT 6
US-09-954-456-1153
; Sequence 1153, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954, 456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233, 617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234, 052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234, 923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235, 134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235, 637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235, 638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235, 711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235, 863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1153
; LENGTH: 10211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1153

Alignment Scores:
Pred. No.: 7,99e-64 Length: 10211
Score: 1139.00 Matches: 703
Percent Similarity: 37.36% Conservative: 556
Best Local Similarity: 20.86% Mismatches: 1052
Query Match: 7,71e Gaps: 1060
DB: 10 Gaps: 139

US-09-150-867-1 (1-2954) x US-09-954-456-1153 (1-10211)
Oy 333 ThrProHisValAsnGlnValLeuAspAspGlnAlaLeuLysArgLysArgLysGlu 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 ACAGAGCTTTCAGAAATTCAGAGCTTGAAGAGACGTTGACAAAGTGAAGAAAGAA 171

```


Db	2182	CAGAAA	---GCTGAGTTCCTCA	-----GATCAGCAACATCAGAG	2217
Ox	960	GIuansngluValThrCysLeuSerGluTyrlYsRheLeuProAsnGluValGluCysLeu			979
Db	2218	GAATATGAAATAATATGCTGTGGAGACCTTCGCAG	---CTTACTGGCGCAAGTTGAAATCTCA	2274	
Ox	980	LysAsnGlnIleSerLysAlaSerGluGlnIleMet			991
Db	2275	GAACACAGAGCTTCAGTTACTCTCAAAATGAATTAATGACAAAGACCGGTGTACCAAGAC	2334		
Ox	992	-----LeuLeuLysGlnGluY			997
Db	2335	TTGCATGCCCAATATATGAGACCTCAGGAGATCTGCTAAATCCAAAGATGCTTCCTGGTG	2394		
Ox	998	-----GluHis	999		
Db	2395	ACAATGAGATCATCATCAGAGAGCTCTTTGGCTTTTGATCAGACAGCCGTGCATGCATCAT	2454		
Ox	1000	Ser---AlaSerIleIleSerLysGlnIleIle	1010		
Db	2455	TCCTTCGCAATATTAATTGGAGACAAAGAGCATGCTTCACAGAGAGTGAATGTCGT	2514		
Ox	1010	-----	1010		
Db	2515	TTAGACAGACCAAGATCCGAAAAATTTGCCATCTCAAAATATAGACTTGATCACTT	2574		
Ox	1011	-----MetGlnGlnSerGluGln	1017		
Db	2575	GAATTTTCATTTAGAGTCTCAAAAACAGATGAACATCAGACCTGCACAAACAGAGTGTAAAG	2634		
Ox	1018	IleLeuGlnLeuThrAspGluValIleThrHisThrGlnSerValGlnGlnThrGluGln	1037		
Db	2635	TTGGTGCAAATCAAGAGCAAAATGAAGAAATCTCATGAAGACAGACAGATGCATCA	2694		
Ox	1038	GlnIleuLeu-----GluMetLysMetHisAspAsp	1048		
Db	2695	AGTTTGTGGCTGAACCAACATCAGCCCATTTATGAATTCACAGAAAGACACATTTGCTCAC	2754		
Ox	1048	-----	1048		
Db	2755	CAGAAATGTTGTGCGAAACCTTAAGTGCCTTGAGAACAGAAAGAAAGACGTGCACAT	2814		
Ox	1049	LeuRheGlnLysTyrlLeuArgAsnLysSerGlu	1059		
Db	2815	TTTAATGATTAAGGTAGAAACTGACGCGCAAGATTTCAAGAAATTAAAAAGACAAACCAT	2874		
Ox	1060	---AlaGluAspLeuLeuArgGluMetGluAsnLeuLysGluThrMetGluSerValGlu	1078		
Db	2875	CTACTTGAACACCTCTCAAAAGAGCATCAACACTTTATCCGAACCCCTA---AGCTTGGAG	2931		
Ox	1079	ValLys-----IleAlaAspThrLysHisGluLeuGlnGluThrIleArg	1093		
Db	2932	AAGAAAGAAATGATGTTCCATCATCTTCTCAAAATTAAGGAAATTAAGAGACGTGACCCA	2991		
Ox	1094	Asp-----LysGlu-----GlnLeuLeuHisGluLysLysTyrlRheRhe	1106		
Db	2992	GAGAAATGGAGCTTTAAGGAAATTAATGCATCTCTTAATCAAGAGAAATGAATCACTTAATC	3051		
Ox	1107	GlnAlaMetGlnThrIleRhe-----ProIleThrProLeu	1118		
Db	3052	CAGAAAAGTGACAGTTTTCGAAACCTTAATATGATGAAGAGGAGAAAGCAATTTTCAGAGTTA	3111		
Ox	1119	SerAspSerLeuProProSerLysLeuVal-----GluGluLysn	1131		
Db	3112	TCCTATCAGTACAGCAAGCAAGAAAACCTATTATTACTCAAAAGATGTGGAAGAAACCGGAAT	3171		
Ox	1132	SerGlnAspProIle-----	1136		
Db	3172	GCATATGAGAGTCTTACTCAAAAATACAAAGACAGCACAGAAAGAAATTTCTAAATTAGAA	3231		
Ox	1136	-----	1136		
Db	3232	TGCTTGCTAAATGAATGACACTAGTCTTTGTGAATAATGAGAAATAAGTGTGGAACAGCTA	3291		

QY	1137	-----GluIleAsnAspTyrHisAsnLeuIleAlaLeuAlaLeuIleGlu	1150
Db	3292	AAGGACGATTTGCAGAAAGAACCCAGAAATTTCTTAACAAAATATGACTTTTGTGAAGAA	3351
QY	1151	ArgAsn---AsnIleMetValCysLeuGluThr-----	1160
Db	3352	AGAAATCAGATCTCGATGCTGATAGAGTTGGAGACAGACGACCAAGCTCTGAGATCTGAGTG	3411
QY	1160	-----	1160
Db	3412	ACAGTAACCAAAACAATTCTAAGAGCGAGCTGGTGTAAAGCAAGAAATCATGACT	3471
QY	1161	-----GluArgAsnSerLeuYsgLugInValIleAspLeu-----AsnThr	1174
Db	3472	TTAAAGGAGAGACAAACAAATGATGCAGAAAGAGATTAACTACTTTTACAGAGAAATGAA	3531
QY	1175	GlnLeu-----GlnSerLeuGlnAlaGlnSerIle	1184
Db	3532	CAGCTGATGAGGATGATGAAGACTAAACATGATCTCAAAATCTGATGAATCGAACCACTAT	3591
QY	1185	GluYsSerAspLeuGln-----LysProGlyGln	1194
Db	3592	AGGAACCTCTGGAAAGAAAGAGAGAGAGAGAAATCAATGTAATTTTAAACCTCAGATG	3651
QY	1195	AspLeuGlnGluGlnGluValLysLeu-----LeuLeuGlnMetGlu	1208
Db	3652	GATCTTGAGATTAAAGAAATTTCTCTAGATAGATTAAATGCGCAGATGGTGCATTAAGAA	3711
QY	1209	LeuLeuYsgLysLeuThrAspSerGlnLeuSerIleGlnLysLeuGlnLeuGlnAsn	1228
Db	3712	GCTATG-----CTAAGAAATTAAGGAATTTAAACTTCAGAGAAAGTAGAGGAAGAGAG	3762
QY	1229	LeuGlnValThrGlnLysLeuGlnThrLeuGlnGlnGlnMetLysAsnIleThrIleGlu	1248
Db	3763	-----GAGTGCTGCGACCATGATTAACAGCAAT-----	3792
QY	1249	ArgAsnGlnLeuGlnThr-----AsnPheGlnAspLeuLysAlaGlnHisAspSer---Leu	1266
Db	3793	AGAGAGAGATCTTGAAACCCACCAATTTTGCAAGACATGCACATGCAGAAATTAATGTCGCTT	3852
QY	1267	Lys-----GlnAspLeuSerGlu	1272
Db	3853	AAAGACTGTGAATATAGATGGCGGAAGAAAGATATTTTCAGGGCCCTCAGTGTGTCACAA	3912
QY	1273	AsnIleGlnGlnSerIleGlnThrGlnAspGlnLeuArgAlaAlaGlnGlnGlnLeuArg	1292
Db	3913	AGTCAAAACGCATGCACACCTTCAGTCTCTCTCCAAACCAACATGACCAACCTGANT	3972
QY	1293	GlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIle	1312
Db	3973	GAGCTAGGAAATATGTGAATATCTGACAGCGCTGGAAGAGTATGAACCTGTAACCTGACCTG	4032
QY	1313	SerSerProAsnHisAspAlaValAlaAsnGlnGlnLysValSer-----	1337
Db	4033	AATGATTCACAGCTCAGAAATGATATCACAGCAACTAGGAAATAAGGCAAGAGAGTAGGAAA	4092
QY	1338	---LeuGlnGlnValAsnSerLeu-----GlnSerGlnMetLeuArgGlyGlnArgAsp	1344
Db	4093	CTACTTAATGAAATTAAATATTAAATGATGACAGTGGCTTCTCATGGTGAAGTTACTG	4152
QY	1345	Glu-----	1345
Db	4153	GAAAGCATACGAGAGGTGAATTTGGTGAACCAACAAATGAACAGCACCCCTGTCTTGG	4212
QY	1345	-----	1345
Db	4213	GCTCATTTGACGAGAGTAAATTCCTACGAGCACTTACATTTGTCAAGCAAAAGATTCAA	4272
QY	1346	-----LeuGlnThrSerCysLysAlaLeu	1353
Db	4273	ATGCACTTTGCCAATTTGCAGAGAAATTTCTATTCTTTAAAGTAGAACACCAATTTTATA	4332

QY 1354 -----ValSerGluLeuGluLeuLeuArgLalaHisValysSer 1366
 Db 4333 CATGATCAGCATCTGATGATGAGCTTAAATCTACAGCTGACAGCCATGTGACTCA 4392
 QY 1367 ValGluGluLeuAsnLeuGluLeuThrLysLysLeuAsnGlyLeuGluLysGluLeu 1386
 Db 4393 TTAAAGCCGAAATTTGGCTTGTCAACGATCTGAAGAACTTTCAGAGTGGTGGT 4452
 QY 1387 GlyLysSerLys-----GluSerGluValLeuLysSerMetLeuGluAsnLeuLys 1403
 Db 4453 AAGAGATGAGCTGGCTGGAGAGAGGGCTGTCATCCCTGTCATCCCTTGTGTG 4512
 QY 1404 GluLysAsnLysLeuLysGluGluLala----- 1414
 Db 4513 CCTACAGCTCTGATCTTACAGAGTTTGGAGACTCCTCTTTACAGAGCTCTTTAGAA 4572
 QY 1415 -----GluLysSerSerLysGluLalaGlnPhe 1423
 Db 4573 CAGACAGAGATATGCTCTTTGATTAATTTAGAGGGCTGTTTCAGCAACAGCTGC 4632
 QY 1424 SerLeuGluGluValPhe----- 1429
 Db 4633 AGTGTAGTGAAGATTTTTCAGAGCATCTGACAGAGAGATCTGACCAGAGAAACCC 4692
 QY 1430 ---SerGlySerGluLysLeuValAspGluLe-----GluValLeuLysAla 1444
 Db 4693 CCTTCGGCCCGCAGAGAGGGTGTGAAGAGCTGATGCTCCTGTGTGGGTGACCGAG 4752
 QY 1445 GlnLeuLysAlaLalaGluArgLeuGlu-----IleLysAspArgAspTyr 1460
 Db 4753 TCCCTCGAAGCTGAAGAGAAATGGAATGCAAGGATTTAGAAATTAAGAAAT 4812
 QY 1461 PheGluLeuValGlnThrAlaAsnThrAsnLeuValGlu---GlyLysLeuGluThrPro 1479
 Db 4813 CAAAGCTCGAGCACTTATTAAGTCTGAAAGCAAGCTGATGCTGCTTACAGACAG 4872
 QY 1480 LeuGlnAlaAspHisGlu-----GluAspSerIleAspArgSerGluGluMetGlu 1497
 Db 4873 TATTTGTCAAGAAATGAACAGTGCACAGACAGAACTGACACGCTGCTGAGATGGAG 4932
 QY 1498 IleLysValLeuGluGluLysLeuGluLysArgAsnGlnTyrLeuLeuGlu----- 1513
 Db 4933 TCCAAAGTTGGCGCAGAAAGAAACAGACAGCAACTGTCACTGAGCTGGAAGTACA 4992
 QY 1514 ArgLeuGlnGluGluLysLeuLeuSerAsnLysLeuGluLalaGluGluMet 1533
 Db 4993 CGACTCCAGCTCAAGAGCTGAGCTTAAGTCTCGG----- 5028
 QY 1534 GluThrSerValLeuLysAspAspLeuGlnGluLysLeuGluSerLeuLeuSerGlu 1553
 Db 5029 -----TCTTGCTGCTGATCGACACAGAAAGATGCTATTCAGGC----- 5067
 QY 1554 AsnIleIleLeuLysGluAsnIleAspThrThrLeuLysHisHisSerAspThrGlnAla 1573
 Db 5068 -----CGAATGAGAGCTGTGACATATCAAAAAGAACATCTCTCAGAACTCAGAA 5118
 QY 1574 GlnLeuGlnLys-----ThrGlnGlnLeuGln 1583
 Db 5119 AGAACACCAAGCATGATGTCATCAGATTGTGATAAGATGTCAGAGAGACCTCAAT 5178
 QY 1584 Leu-----AlaLysAsnLeuAlaIleAlaIleAspAspAsnProIle 1598
 Db 5179 CTAGACATGAGAAATACTGAGACTGTGCTTGAACCCACAGAGAGAGCTCTGGG 5238
 QY 1599 ThrGlnGluLysGluThr-----SerAla 1606
 Db 5239 GAACAGTCCCGACATACCAATTTAGCCTCCAGGGGAGATAAACCCAGGGCTCTTCA 5298
 QY 1607 AspCysValHisProLeu----- 1612
 Db 5299 GAATGATTTCTGAATTTGATTTCTGTCTTAATGCTTTGGTACCTATGATTTCTTG 5358
 QY 1613 -----GluGluLysIleLeuLeuLeuThrGluGluLeuHisGlnLysThrAsnGluGln 1630

Db 5359 GGGATTCGGAGATATCCATTAATCTTCAACTGGCGGTAAAGAGACATCAATGGAAT 5418
 QY 1631 GluLysLeuLeuHisGluLysAsnGluLeuGluGlnAlaGlnValGluLeuLysCysGlu 1650
 Db 5419 TTGATATTACTTCAATGTGATAGAGACCGTGACAGA-----AAA 5457
 QY 1651 ValGluHisLeuMetLysSerMetIleGluSerLysSerSerLeuGluSerLeuGlnHis 1670
 Db 5458 GTTGAAGTTTGCTAAATGAAATGAATTAACATCTCAAACTC----- 5502
 QY 1671 GluLysHisAspThrGlnGlnGlnLeuLala-----LeuLys 1683
 Db 5503 -----CATTTACAGAGAGTCAACTAATGACCAAAATTGAAGCATCAATGAA 5556
 QY 1684 GlnGluMetGlnValValThrGlnGluLysGluLeuGlnGlnIleHisGlnHisLeu 1703
 Db 5557 AAAATAGTTGGGGAAGCTTACAGAAAGAACTGAGTTTAAGTGAATAATGGAATATTT 5616
 QY 1704 ThrAlaGluValAspHisLeuLysGluAsnIleGluLeu-----GlyLeuAsnPheLys 1721
 Db 5617 TCTTGATACACAGAGATTAOTCCAGAGATGAAGAACTCTGAAGGCTCAATTCAT 5676
 QY 1722 AsnGluAlaGln---GlnLysThrThrLysGluGlnCysLeuLeuAsnGluAsnLys 1739
 Db 5677 TTGAATATCATCAGATTAATCATCAGCTGAAGATTTGGAGATATGTGCGCAAGTGT 5736
 QY 1740 -----GluLeuGluGlnSerGlnHisArgLeuGlnCys 1750
 Db 5737 AATGACAGCTGAGAGAGAGATTTCTGTATGTGAAATGAGACTGAGATGATGAGATCG 5796
 QY 1751 Glu-----IleGluGluLeuMetLysSerLeuLys 1760
 Db 5797 GAGAAAGCTAGCATGATGATGATGAGCCCTACCTACGGAGCTGACTGAGATGACTTAA 5856
 QY 1761 AspLysGluSerAlaLeuGlnThrLeuLysGluSerGluGlnLysValIle---AsnLeu 1779
 Db 5857 ACAGAGAGCTATGTTTAAAGAAAGCAATGAATAAGCACAGATTAATGTCTGCTT 5916
 QY 1780 AsnGlnGluMetGluMetValMetLeuGluGluLeuLysAsnSerGlnArgThr 1799
 Db 5917 GAGAGAAAGCTCAGCTGTCTACAGAGAGAGAAACACCTTCGTGGAGAAATTAGATCT 5976
 QY 1800 ValIleAlaGluArg-----AspGlnLeuGlnAspLeuArgLysValGlu 1816
 Db 5977 ATGTCAAAAACCAACGCGCATGTGATCTGTCTGAAAAATGAGAGAGAAACACA 6036
 QY 1817 MetSerIleGluThr-----GlnAspAsp 1824
 Db 6037 ---GAGCTTGAAGTCTATCAAGTGAAGTGTCTCATTCATTCAGGTGCGACAGCGACAG 6093
 QY 1825 LeuArgLysAlaGlnGlnAlaLeuGlnGlnLysAspLysValGlnGluLeuThrSer 1844
 Db 6094 GTGAAGGAAAGAGAGCAACTCTTCAGACTTGTCTCTGATGTAGTGAAGCTTTAA 6153
 QY 1845 GlnIleSerValLeuGlnGlnLysIleSerLeuGlnGluAsnGlnMetLeuTyrAsnVal 1864
 Db 6154 GACAAACTCATCTCCAGAAAGCTCAGAGATTGTGAAAGAGACTCACAGGCACTGTCT 6213
 QY 1865 AlaThrValLysGluThrLeuSerGluArgAspAspLeuAsnIleSerLysGlnHisLeu 1884
 Db 6214 TTGACAAATGTGAGTGAAGAAACCAATTCACAACTGAATTAAGAGAAAGAAATTCCT 6273
 QY 1885 PheSerGluIleGluThrLeuSerLeuSerLeuLysGluLysGluPhe----- 1900
 Db 6274 GTCAAGATGTGAAGACCTTCAGGCGCAGACTGATGATATCAGATTAAGAAAGCTGAAT 6333
 QY 1901 -----AlaLeuGluGlnAlaGluLysAspLysAlaAspAlaAlaArgLysThrIle 1917
 Db 6334 GTCTCAAGGCTTGAAGCGGCACTGTGTGAGAAAGTGAAGTTCGATTTGAGGCTGAGC 6393
 QY 1918 AspIleThrGluLysIleSerAsnIle----- 1926

Db 6394 TCACACAGAGAGATGCATCAGTCAGAGAGCATGCAGAACTGAGATTGCAT 6453
 QY 1927 -----GluGluGlnLeuLeuGlnGlnIleThrAsnLeuLysGluThrLeuTyrGlu 1943
 Db 6454 GAGCCGATGTAAGAAAGAACCTGCACATCGCAGAGAAACGAAAGACGGAGCGGAG 6513
 QY 1944 ArgGluSerLeuIleGlnCysLysGlnLeuAlaLeuAsnThrGlnIleLeuTyrGlu 1963
 Db 6514 AATGATTCACCT-----AAGGATTAAGTTGGAACCTTGAAGGAAATGCAAGT 6564
 QY 1964 ThrLeuLysSerLysAspLeuAlaLeuGlnLysMetGluGlnIleTyrAspGlnAla 1983
 Db 6565 TCACAGAGAAACACAGAGACTGATGAT-----CTTGATGCCGCAAAATCCAAACCA--- 6615
 QY 1984 AsnLysValIleAlaLeuThrGlnLysMetSerSerLeuGlnGlnIleAsnGlnAsn 2003
 Db 6616 -----GAACTAGACACTTAATAACCAAAATGCAAGATG 6651
 QY 2004 ValThrThrLeuLysGluGlnGlu-----GlyGlnLysGluThr 2016
 Db 6652 GCCAGAGCCCTGAAAGTTTGAATTAACCTTGCACGTTAAGGCTGAAAGAAAT 6711
 QY 2017 PheTyrLeuGlnArgProSerLysGlnIleSerSerGlnMetGluLeuArgGlu 2036
 Db 6712 CTGACAAAACAAATACAGAAACAAACAGCTGATGACAGACTGACAGACTGATCTCT 6771
 QY 2037 SerLeuLysThrLysAspLeuGlnLeuGlnGlnIleGlnLysGlnIleSerGlnAlaThr 2056
 Db 6772 TCATTTAAAGCTCTTGAAGAAAGAGCAAGCAGATACAGATTAAGAAAGAACT 6831
 QY 2057 -----AsnGlnLysAsnLeuThrAlaLysIleSerSer 2068
 Db 6832 AAAATCGATGAGATGCTCAGATACAGTTAAAGAGCTTAATGAGAGCTAGCAGAC 6891
 QY 2069 Leu-----GluGlnGluIleLeuGln---AsnAlaSerIleLeuAsnGlnAlaValSer 2085
 Db 6892 TTGCTGTGTCACCAAGAAATATGATGAGCCACAGAACAGCTGACCCACCATAGAG 6951
 QY 2086 GluTyrGlnAsnLeuArgGlnSerLysGlnIleLeuValSerGlnLeuGln--- 2102
 Db 6952 GAAAGCATCGCTGAGAAATAGCATTTGAAAGCTGAGAGCCGCTGAAAGCTGATGAA 7011
 QY 2103 -----GlnLeuSerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArg 2120
 Db 7012 AAGAGCAGCTGCTGCTGCTTCAACAACTGAG-----GAAAGTAGCAT 7056
 QY 2121 GluLysAspGlnAlaValAsnLysIleAlaSerLeuAlaGlnIleLysIleLeuThr 2140
 Db 7057 CATCCAGATTACTTAAGGTAGAGTGAGAACTTGAAGAGAGCTGAGATAGCCAGG 7116
 QY 2141 LysGlnMetAsp-----GluPheArgAspSerLys-----GluSerLeu 2153
 Db 7117 ACAAAACCAAGACATGCAGCTCTTGAGGCAGAGAAATTCAAAGAGAGCTGAGAGCCCTA 7176
 QY 2154 GlnGluGlnSerSerHisLeuSerSerGlnIleLeuCysThrTyrLysThrGlnLeuGlnMet 2173
 Db 7177 AAACCAAAATAGAGAGGATGACCCAAAGCTGAGAGGCTGGAATTAAGATTGTTACT 7236
 QY 2174 LeuLysGlnGlnLysGlnAspIleAsnAsnLysLeuAlaGlnLysValLysGlnValAsp 2193
 Db 7237 ATAAGGTCAGAGAAAGAAATCTGACAAATGAATTACAAAGAGCAAGCATGATCT 7296
 QY 2194 GluLeuLeuGlnHisLeuSerSerLeuLysGlnIleLeuAspGlnIleGlnMetGlnLeu 2213
 Db 7297 GAATTA-----GAAATA 7308
 QY 2214 ArgAsnGlnLysLeuArgAsnTyrGlnLeuCysGlnLysMetAspIleMetGlnLysGln 2233
 Db 7309 AATAATCATCATTTGAAAT----- 7329
 QY 2234 IleSerValIleLeuArgLeuMetGlnAsnGlnProGlnGlnGlnIleAspAspValAlaGlu 2253
 Db 7330 -----ATTGTGCAAGAAAGAGCAAGAGAAAGTACAGATGAAGAAATCAAGCACT 7383

QY 2254 ArgMetAspIleLeuGluSerArgAsnGlnGlnIleGlnIleLeuMetGlnLysIleSer 2273
 Db 7384 GCCATGAGATGCTTCAACA-----CAATTAAGAGCTCAATAGAGAGTGGCA 7434
 QY 2274 AlaValTyrSerGlnGlnHisThrLeuLeuSerSerLeuSerSerGlnLeuGlnLysGlu 2293
 Db 7435 GCCCTGCATTAATGACCA----- 7452
 QY 2294 ThrGlnAlaHisLysHisCysMetLeuAsnIleLysGluSerLeuSerSerThrLeuSer 2313
 Db 7453 ---GAAGCCTGAG-----GCCAAAGACAGAACTTCTAGT 7485
 QY 2314 ArgSerPheGlySerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeu 2333
 Db 7486 AGTCAAGTAGAGTCTGTAACCTTAGAAGGCTCAGTTGCTTCAAGAGCCCTGATAGGCC 7545
 QY 2334 LeuAsnLysPheLysValValTyrArgThrAlaAlaValLysGlnAspHisSerLeuIle 2353
 Db 7546 AAAATTAATTAATTAATGTTTGCATCTCAGTG-----AATGCCCTCAT 7590
 QY 2354 LysAspTyrGlnLysAspLeuAlaAlaGlnGlnLysArgHisAspGlu-----Leu 2370
 Db 7591 CAAGAGTAGAAGATGCGACAGCAAGCAAGCTGAGAAAGAGATGAAGAAATCAGTAGACTG 7650
 QY 2371 ArgLeuGlnLeuGlnCysLeuGlnGlnHisGlyArgLysTyrPheSerAspSerAlaSerGlu 2390
 Db 7651 AAAATCAAAATTCAGAACCCAGACAGCAGCTTGTCTTAACGTCCACAGCTGAGAGAGAG 7710
 QY 2391 GluLeuLysPhe-----CysGlnIleGlnPheLeuAsnGlnLeuLeuPheLys 2406
 Db 7711 CACCAACTTGGAGAGACCAAAATCTAGAACATGAGAAATTCAGAGTGAATGGAGCAG 7770
 QY 2407 LysAlaAsnIleIleGlnSerValGlnAspPheSerGlnValGlnValPheLeuAsn 2426
 Db 7771 AAGATCAAGGCTCAATCCAAATCCCTTTGACAGACACATTAGAACTGCTGACG 7830
 QY 2427 GlnValGlySerThrLeuGlnGlnLeuGln---HisLysLysGlyPheMetGlnTyr 2445
 Db 7831 AGTTTTCAGAAATCTGAGAAATGAGCTTCAATTAATGACAAATAGCAAAATGCTCTT 7890
 QY 2446 LeuGlnGlnPheGly-----AspLeuHisVal 2454
 Db 7891 GTTGAATAAGTAACAAATGACTCAAGAGAACTGAGTCAGAGGAAATGCAATGAG 7950
 QY 2455 AspAlaLysLys-----LeuSerGlnGlnMetGlnGlnIleAsnArgArgIleAla 2471
 Db 7951 ATGCACAGAAACAGCAGAGCTGTCAGAAAGAACTGAGAGAAATAGCTTACT 8010
 QY 2472 SerThrIleGlnLeuThrLysArgLeuLysAlaValAlaGlnSerLysIleGlnArg 2491
 Db 8011 GGAGAGTTCAGTTCCTTGTGGAATTAAGAC-----AGCAAAAGATCA--- 8058
 QY 2492 GluIleThrValTyrLeuAsnGlnPheGlnAlaLysLeuGlnGlnLysLysGlnIleAsn 2511
 Db 8059 -----TTGAAGAGCTCACACTAGAAAT 8082
 QY 2512 LysGlnLeuMetArgArgMetGlnHisHisGlyProSerAlaSerValIleMetGlnGln 2531
 Db 8083 AGTGATTAAGAGAG-----ACCTAGATTCATGATCACAAGAAC 8121
 QY 2532 AsnAlaArgLeuLeuGlyIleLeuLysThrValGlnAspGlnSerLysLeuGlnSer 2551
 Db 8122 CAGGTGAGAAAGAGAGG-----AAAGTAGAGAGAAATAGCTGAATATCAGCTA 8172
 QY 2552 ArgIleLysMetLeuGlnAsnGlnLeuAsnLeuValLysAspAspAlaMetHisLysGly 2571
 Db 8173 CGG-----CTTCATGAACT 8187
 QY 2572 GluLys-----ValAlaIleLeuGlnAspLysLeuLeuSerArgAsnAlaGlnAlaGlu 2589
 Db 8188 GAAAGAGAAACACAGGCTTGTCTTTGGAC-----ACAAACAAAGCTATGAAAGTAGAA 8241

QY 2590 LeuAsnAlaMetGlnValLysLeuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLys 2609
 Db 8242 ATCCAGACATACCGAGAAATGACTTTAAAGAAATGCTCAGTTCACAGAAAGCTG 8301
 QY 2610 GlnIleuLysLeuGlnLysMetValAlaLysGlyAlaValProLysLysGlnLys 2629
 Db 8302 GAGATAGACCTTTAAAGTCTACT-----AAAGAAAGCTC 8337
 QY 2630 AspAsn-----LeuLysThrLysValLysLysIleGlnMetGln 2642
 Db 8338 AATATATCATTTGAAGTACTACTACATTTTGGAGAAATTTGAAGAAACCAAGATGAG 8397
 QY 2643 LysIleLysLys-----SerLysAlaThrAspGlnGlnLys 2654
 Db 8398 AATCTAAATATGTAAATCAGTTGAGAGAAATGACGTGCCAGGGGAAATGAG 8457
 QY 2655 AlaThrLysLysSerCys-----LeuGlnAspLysGlnLysLysLysLysLys 2672
 Db 8458 TTGTTGATCAATCTCTGTAACACGTGAGAGAAAGAGATCTGCAGAAAGAACTC 8517
 QY 2673 GlnLysLeuArgArgAlaGlnAlaAspAsnAspThrValCysValProLysAspPyr 2692
 Db 8518 TCTCAACTTCAAGCTGCGACAGAGAGAGAAACA----- 8553
 QY 2693 GlnLysAlaSerThrPheProValThrCysGlyGlySerGlyIleValGlnSerThr 2712
 Db 8554 -----GGTACTGTTATGATACC 8571
 QY 2713 AlaMetLeuValLeuGlnSerGln-----LysAlaAlaLeuGln 2725
 Db 8572 AAGCTCGATGATTTAACTAGATCAAGAACTGAAAGAACTGTAAGAAACCA 8631
 QY 2726 ArgGlnLeuSerHisThrLysLysLysThrHisLysLeuSerArgThrMetSerSer 2745
 Db 8632 AAGGAGGACATGAATACTTGGATTAAGTACTGTTCTCTG-----CTATTAAGCAT 8662
 QY 2746 GlnAspArgLysLysThrLysAlaLysSerAspAlaHisSerHisThrGlySerSer 2765
 Db 8663 GAAAGTTAGAGAAAGCAAGAGATGTTAGACACAGTGGCCATCTGTTCACAG 8742
 QY 2766 His-----ArgLysSerProHisLysThrGlnThrArgHisGlyPro 2780
 Db 8743 CAATCTAAACAGATTCCTCCAGAGGCTCTCT-----TTCCTAGGTCA 8784
 QY 2781 ValThrProGlnArgSerGlnMetProSerLeu-----HisLeuGlySerProLys 2797
 Db 8785 GTTGTTCAGACATCTCCATCTCTCTCTGTACTGAAAGAGTGTATCATCTGGCCA 8844
 QY 2798 LysSerGlnSerSerThrLysArgValLysSer----- 2808
 Db 8845 AATTAAGCTTCAGCGAAGAGCGCAAGATCCAGTGAATATGGAGAAATGGTGAGACCA 8904
 QY 2809 -----ProAsnArgSerGlnLysLysSer-----GlnLeuValMetSer 2821
 Db 8905 AACCTGCTACCCCGACAGAGCTTTCTTAAAAAACAGAAAGCATGATGATGTAT 8964
 QY 2822 -----ProGlnLys----- 2824
 Db 8965 CACCTCGAGAGACAGAGAGTACTGAGTTCAGCCAGAGGAGCTTCAGAAAGTTGA 9024
 QY 2825 -----ThrGlyMetHisLysHisLys----- 2831
 Db 9025 AAGAAAGGGTTGCTGACATCCCGACAGAGAAAGTACGCCCATATATCTCTCGAAAGACA 9084
 QY 2832 -----LeuSerProSerLysValGlyLeuHisLysLysArgAlaLeuSerPro 2847
 Db 9085 ACCATGGCACTCGACACAGCCCCCTGGCTGGCAGAG-AAAGTTAGGCTTATCCCACT 9143
 QY 2848 -----AsnArgSe 2850
 Db 9144 GAGTCTCGGAAAGAAATCTTGCAGAGTCTCCAAACCAAGCTGTGGAGAGCATGC 9203
 QY 2850 rgluMetProThrGlnHisValLysSerProGlyLysThrGly-----LeuHisLysAs 2868

Db 9204 ACAAAGGCTAAAGTTGCTCAGCGAGGCCAGTAGATTCAGGACACATCTCCAGAAACC 9263
 QY 2868 nLeuThrGlnSerThrLysPheAspAsnLeuSerSerProCysLysGlnLysValGln 2888
 Db 9264 CACCAAGAAATCGTCCAGTCAATATCTTCGAGAGAAAGTCCGACTGCAGCCCCAG 9323
 QY 2888 nGlnAsnLeuAsnSerProLysGlyLysLeuPhe-----AspValLysSe 2903
 Db 9324 AAGAGGCTCGAGGCTCAGCGAGGCCGACTGTCTCCAGCCCCAAAGCTGGACGTGACTC 9383
 QY 2903 rLysSerMetProLysCysProSerGln 2912
 Db 9384 CAAGGCGACTGAGAACTGTAAAGTCCAG 9411
 RESULT 7
 US-09-967-768A-186
 / Sequence 186, Application US/09967768A
 / Patent No. US20020150877A1
 / GENERAL INFORMATION:
 / APPLICANT: Augustus, Meena
 / TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sigm
 / TITLE OF INVENTION: Sets
 / FILE REFERENCE: 689290-72
 / CURRENT APPLICATION NUMBER: US/09/967,768A
 / PRIOR FILING DATE: 2001-09-28
 / PRIOR APPLICATION NUMBER: US/60/236,109
 / PRIOR FILING DATE: 2000-09-28
 / PRIOR APPLICATION NUMBER: US/60/236,034
 / PRIOR FILING DATE: 2000-09-28
 / PRIOR APPLICATION NUMBER: US/60/236,111
 / NUMBER OF SEQ ID NOS: 325
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 186
 / LENGTH: 10211
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / US-09-967-768A-186
 Alignment Scores:
 Pred. No.: 7,99e-64 Length: 10211
 Score: 1139.00 Matches: 703
 Percent Similarity: 37.36% Conservative: 556
 Best Local Similarity: 20.86% Mismatches: 1052
 Query Match: 7,71% Indels: 1060
 DB: 10 Gaps: 139
 US-09-150-867-1 (1-2954) x US-09-967-768A-186 (1-10211)
 QY 333 ThrProHisValAsnGlnValLeuAspAspGlnAlaLeuLeuLysArgLysLysGln 352
 Db 112 ACAAGAGCTCTTCAGAAATTCAGAGCTTGAAGAGACGTTGACAACTGAGAAAGAA 171
 QY 353 IleLeuAspLeuLysGlnLeuGlnAsnLeuGlnLysSerSerGlnThrLysAlaGln 372
 Db 172 AAGCAGCAAGAGCGAGTTCAGCTGACAGTCTGAGGCTCGCCCTCAGAGCAACAG 231
 QY 373 AlaMetAlaLysGlnLys-----HisThrGlnLeuAla 384
 Db 232 AAGCTTGAATATGAAAGAAACCGAGGTTACAAACCTGAAAGAGGAGAAATCAAGATTGATG 291
 QY 385 GlnLysLysGlnLeuHisLysGlnArgGlnAspArgLysLysLeuThrHisLeuThrAsnIleVal 404
 Db 292 GAATATATGTAAGTCTCGAGAAACCTGAAGACAGATTTCTCAT-----GAACTTCAA 345
 QY 405 ValAlaSerSerGlnLysSerGlnLysPhe-----GlnArg 416
 Db 346 CTCAGAGAGTCAACAATTTCCAGAGAGGACCACTGAATTCAGGCAAAACCAATATA 405
 QY 417 ValLysArgLysArgArgValThrThrAlaProLysLysLysGlnAsnSerLeuHisAla 436
 Db 406 GAAAAAGCTGGAACAGAACTTAAAGGTGTAATCTAGCTTGAAGAGGCAACAAAGCT 465

Oy 437 SerGlyValSerAspPheAspMet-----Leu 445
 Db 466 GCGCAGTGTGCGATGCTCTCTGATCCATGCAATACACCAAAAAATTTTACAACT 525
 Oy 446 SerArgLeuProGlyAsnPheSerLysAlaLysPheSerAspMetProSer----- 463
 Db 526 CCACATACACCAAGTCATATTATAGTGTCCAGGTATGAAGACTTAAGAAAAATAT 585
 Oy 464 PheProGluIleAspAspSer-----ValCysThrGluPheSer 476
 Db 586 AATAAAGAGGTGAAGAAAGAAAAAGATTAGAGCAGAGGTTAAAGCCTTGAGCCTAA 645
 Oy 477 AspPheAspAlaLeuSerMetMetAspSerAsnGlyIleAspAlaGluIleProAsnLeu 496
 Db 646 AAAGCAAGCCAGACTCTTCCACAGCCACCATGAATCACCGGACATTTGCCGCAATCAG 705
 Oy 497 AlaSerLysValThr-----HisArgGlyLysThrSerLeuHis----- 509
 Db 706 GCTTCATCATCTGTGTTCTCATGCGACAGAGAAAGACCCCAAGTCATCTTTCATCTAAT 765
 Oy 509 ----- 509
 Db 766 TCTCAAGAACCTCCAAATTAGAGAGATTCTCTGCATCTTACTTTCTGGGAAACAAG 825
 Oy 510 -----GlnSerMetIleAspPheGlyGlnIleSerAspSerValGlnPheHis 825
 Db 826 GTGACTCCAAAGTCCATCACTTTCGAAATAGGAAAAAGAGATGCTTAATAGCAGTTTCTTT 885
 Oy 526 AspSerSerLysGluAsnGlnLeuGlnTyrLeuProLysAspSerGlyAspMetAlaGlu 545
 Db 886 GACATTTCTAGCAGTCTCATCTTTTG-----GATCAATATAAA 924
 Oy 546 CysArgLysAlaSerPheGluLysGluIleThrSerLeuGlnGlnLeuGlnSerLys 565
 Db 925 GCGGAGAAATCAAGAGCTAAGAAACAGATTAATGAGTTGGAACCTGCGCTGCAAGAGACAT 984
 Oy 566 GlnGlnGlnLysLysGlnLeuValGlnSerPheGluLeuLysIleAlaGlnLeuGlnGlu 585
 Db 985 GAAAAAGAAATGAAGAGCCAGTGAATAAGTTTCA-----GAACTCCAACATC 1032
 Oy 586 GlnLeuSerValLysAlaLys-----AsnLeuGlnMetValThrAsnSer--- 600
 Db 1033 CAATCTGAG-----AAAGCAAAAGTGGAATTAATTGAAAAAGAAATTTTGAACAAATGT 1089
 Oy 601 ArgGlnHisSerIleAsnAlaGluValGlnThrAspValGluLysGluValAlaArgLys 620
 Db 1090 AGGATGAACTAGTGAGAACACACACAAATACGAC----- 1125
 Oy 621 GlnMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSer 640
 Db 1126 -----CAGGCTCAACCAAGTATACTGCATTTGGAACAAAACTGAAAAAATG 1173
 Oy 641 SerValAspGlyLysArgLeuSerSerSerHisAspGlyCysIleGlnHisArgLysMet 660
 Db 1174 ACGGCAAGAT-----TTGAGTTGTGACGACAAATGCAAGAAAGTGGCAGAGTTCT 1224
 Oy 661 LeuGlnGlnLysIleValAspLeuGln-----GluPheIleGluAsnLeuAsnLysLysSer 679
 Db 1225 CTGGAACAGAAATTAAGAAAAAGAAAGAGATTCAAGAGAGACTCTCCGT----- 1278
 Oy 680 GlnAsnAspLysGlnLysSerSerGlnAsnAspPheMetGluSerIleGlnLeuGlnGlu 699
 Db 1279 -----CAACAGCGTCTTCTCCAAACACTGACCCAGAGAGTGCATCAGATG----- 1323
 Oy 700 AlaIleMetAlaGluLysAlaAsnAlaLeuGlnGluLeuAlaLeuMetArgAspAsnPhe 719
 Db 1324 -----AAGGCCAGACTCACCCAGAGATTACAGCAAGCAAG-----AATATG 1365
 Oy 720 AspAsnIleIle-----LeuGlnAsn 726
 Db 1366 CACAACGCTCTGACAGGCTGAAGTGAATAACTCAGTATGAAGAACACAGCTTGAAGAAC 1425

Oy 727 -----GluThrLeuLysArgGluIleAlaAspLeuGluArgSer----- 739
 Db 1426 AATTGGAAAGACTTTTAGCAAAAGTTTGTCACAGCTGAACAGGCTTCCAGCCGAGTCAG 1485
 Oy 740 LeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGlnLysGluThrGln----- 757
 Db 1486 ATCAAGAGAAATAGACTGAGAGAGACATGAGAAATGAAAGAAAGAAAAACAACTCCTT 1545
 Oy 758 LysGlnHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsn 777
 Db 1546 AAGAGTCACTGACCAAAAGGCCAGAGAAAGTCTCCACCTTGAGACAGCAACTCAAGAAC 1605
 Oy 778 AlaGlnMet-----TyrAsnGlnAsnLeuGlnGluAspLeuGlnThrLysThrLys 794
 Db 1606 ATCAACAGCTGTTTAATCAGAGCCAGAAATTTGCAAGAAAGAAATGAACCGAAGAT--- 1662
 Oy 795 LeuLeuLysGlnGlnGluIleGlnLeuAlaGluLeuArgLysArgAlaAspAsnLeuGln 814
 Db 1663 -----ACCTTCAGAGAAACCATGTTAAGAGATCTTCAAGAAAAATTAATCAGCAGAA 1716
 Oy 815 LysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysLeu----- 832
 Db 1717 AACTCCTTGACTTTAGAAAAAACTGAAGCTGCTGCTGATCTGGAAGAACCGCAGAT 1776
 Oy 833 CysGlnGlu-----IlePheGlnLeuLysGlnSerLeu 843
 Db 1777 TCTTCTCAAGACCTTTTGAAGAAAAAGACAACTCACATTTGAACAACTTATATGATTA 1836
 Oy 844 SerAspAlaGluAlaValThrArgAspAlaGlnLysGluCysSerPheLeuArgSerGlu 863
 Db 1837 AGCAAGACAGAGAAAGAGATCCAA-----GCCCTGCGAGTCTTGAAGTTAAAAAGAA 1893
 Oy 864 AsnLeuGlnLeuLysGlnLysMetGluAspThrSerAsnTyrPyrAsnGlnLysGluLys 883
 Db 1894 GAAATGTGAATGAAGAAAGAAAACTCTGTTTCTTGTGAAAGTGAAGAAAAAGAAAA 1953
 Oy 884 AlaAlaSerLeuPheGluLysGlnLeuGlnTyrGlnLysSerAsn----- 898
 Db 1954 CTTTAAC-----CAGATGAAATCAGAAAAAGAAAACTTCGACAGTAAAT 2001
 Oy 899 -----TyrLysLysMet 902
 Db 2002 AATCACTTGAATCTGTCTGAAGACACAGCAAAATTAAGATCAATACAAACAGAGA 2061
 Oy 903 GluAlaAspLeuGlnLysGlnLeuGlnSerAlaPheAsnGluIleAsnTyrLeuAsnGly 922
 Db 2062 GTAAGACGCTGAGATGAGACAGAGAAACCTTAAGTGTGAGATCAGAAACCTTCAACAA 2121
 Oy 923 LeuLeuAlaGlyLys-----ValProArgAspLeuLeuSerArgValGlnLeuGln 939
 Db 2122 GTGTTAGACAGTAACTCAGTGAAGTAGAGACCCAGAAACTGATTATATGAGAGCTACAG 2181
 Oy 940 LysLysValSerGluPheSerLysGlnLeuGlnLysAlaLeuGlnGluLysAsnAlaLeu 959
 Db 2182 CAGAAA-----GCTGAGTTCTCA-----GATCAGAAACATCAGAAAG 2217
 Oy 960 GlnAsnGluValThrCysLeuSerGluTyrLysPheLeuProAsnGluValGluCysLeu 979
 Db 2218 GAAATAGAAATATATGTGTTGAAGACTTCTCG-----CTTACTGGCAAGTTGAATCTTA 2274
 Oy 980 LysAsnGlnIleSerLysAlaSerGlnGluIleMet----- 991
 Db 2275 GAACACAGCTTCACTTACTGTCAAATGAATTAATGAACAAAGACCGGTTACCAAGAC 2334
 Oy 992 -----LeuLeuLysGlnGlnGly----- 997
 Db 2335 TTGATGCCGATATATAGAGAGCTCAGAGATCTGCTAATAATCAAGATGCTTCTGCTG 2394
 Oy 998 -----GluHis 999
 Db 2395 ACAATGAAGATCATCAGAGAACTTTGGCTTTTGATTCAGACCGCTGCATGCATCAT 2454
 Oy 1000 Ser-----AlaSerIleIleSerLysGlnGluIleIle----- 1010

Db 2455 TCCTTGCATAATATTAATGAGACAGAGCATGCTTCAGAGAGAGTGAATGCTCGT 2514
 QY 1010 -----LysProlyGln 1194
 Db 2515 TTAGAAGCAGACCAAAATGCGAAAAATTCGCATCTCTCAAAAATAGAGTTGATTCACCT 2574
 QY 1011 -----MetGlnGlnGlnSerGlnGln 1017
 Db 2575 GAATTTTCATTAGAGTCTCAAAAAACAGATGACACTGACAGCTCGCAAAAGCATGTGAGAG 2634
 QY 1018 ILLeuGlnLeuThrAspGluValThrHisThrGlnSerLysValGlnGlnThrGlnGlu 1037
 Db 2635 TTGGTGCATAATCAAGAGAGAAATAGAGAAATCTCATGAAACAGACAGCATGATCA 2694
 QY 1038 GlnThrLeu-----GlnMetLysMetHisAsp 1048
 Db 2695 AGTTTGTGCTGAAACAGATCAGCCGATTAAGTTACAGAGAGACACTTCTGCTCAC 2754
 QY 1048 ----- 1048
 Db 2755 CAGATCTGTGCTGAACCTTAAGTCCCTTGAGACAGAGAAAAAGCTGCAACT 2814
 QY 1049 LeuPheGlnLysTyrIleArgAsnLysSerGlu----- 1059
 Db 2815 TTTAAATGATTAAGTAGAAGACTGACAGCAGAGATTCAGAAATTAAGAGCAACCAT 2874
 QY 1060 ---AlaGlnAspLeuLeuArgGlnMetGlnAsnLeuLysGlnThrMetGlnSerValGlu 1078
 Db 2875 CTACTTGAAGCTCTTAAGAGAGCTTAACTTATCCGAAACCTTA---AGCTTGAG 2931
 QY 1079 ValLys-----IleAlaAspThrLysHisGlnLeuGlnGlnThrIleArg 1093
 Db 2932 AAGAAAGAAATGAGATTCCTCATCATTTCTCTAAATTAAGGAAATTAAGAGCTGACCAA 2991
 QY 1094 Asp-----LysGln-----GlnLeuLeuHisGlnLysLysTyrPhePhe 1106
 Db 2992 GAGATGAGGACTCTTAAGAAATTAATGCAATCTTAATCAAGAGAGATGCAACTATC 3051
 QY 1107 GlnAlaMetGlnThrIlePhe-----ProIleThrProLeu 1118
 Db 3052 CAGAAAGAGTGAGAGTTTGCAAACTATATATGATGAAGAGGAGAAAGACATTTTCAGACTTA 3111
 QY 1119 SerAspSerLeuProProSerLysLeuVal-----GlnGlyAsn 1131
 Db 3112 TCTGATGAGTACAGCAAGAAATTAATTTACTACAAAGATGTGAAGAAACCGCAAT 3171
 QY 1132 SerGlnAspProIle----- 1136
 Db 3172 GCATATGAGATCTTAAGTCAAAAATACAAAGCAGACAGAAAGAAATTTAAATTAGAA 3231
 QY 1136 ----- 1136
 Db 3232 TGCCTTGCATAATGATGCATAGCTTTGTGAATAATAGAAATAGATTGGAAAGCTA 3291
 QY 1137 -----GlnIleAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGlu 1150
 Db 3292 AAGGAGCATTTTGCAAAGAGAACACCAAGAAATTTTACAAATATAGCATTTGCTGAAGAA 3351
 QY 1151 ArgAsn---AsnIleMetValCysLeuGlnThr----- 1160
 Db 3352 AGAAATCAGAAATCTGATCTAGAGTTGAGACAGTGCAGCAAGCTCTGAGATCTGAGATG 3411
 QY 1160 ----- 1160
 Db 3412 ACAGATTAACCAAAATTTCTAAGAGCAGAGCTGTGTTAAAGCAAGAAATCATGACT 3471
 QY 1161 -----GluArgAsnSerLeuLysGlnValIleAspLeu-----AsnThr 1174
 Db 3472 TTTAAAGAGAAACAAACAAATATGCAAAAGAGTTAATGACTTTATTCAGAGAGATGAA 3531
 QY 1175 GlnLeu-----GlnSerLeuGlnAlaGlnSerIle 1184

Db 3532 CAGCTGATGAGGTATATGAAAGACTAAACATGAAATGTCAAAATCTAGAAATCAGAACCAATT 3591
 QY 1185 GlnLysSerAspLeuGln-----LysProlyGln 1194
 Db 3592 AGGAACCTCTGTGAAGAAAGAGAGAGTGAGAGAAATCAATGTAATTTTAAACCTCAGATG 3651
 QY 1195 AspLeuGlnGlnGlnValValLysLeu-----LeuLeuGlnMetGln 1208
 Db 3652 GATCTTGAAGTTAAGAAATTTCTAGATAGTTAATATGCCAGTGTGTCGAATTAGAA 3711
 QY 1209 LeuLeuLysGlnHisLeuThrAspSerGlnLeuSerIleGlnLysLeuGlnLeuLysn 1228
 Db 3712 GCTATG-----CTAAGAAATTAAGGATTAATAACTTCAGAGAAAGTGAAGAGAGAG 3762
 QY 1229 LeuGlnValThrGlnLysLeuGlnThrLeuGlnGlnGlnMetLysAsnIleThrIleGln 1248
 Db 3763 -----GAGTGCCCGCAGCATGATTAACAGCAAT 3792
 QY 1249 ArgAsnGlnLeuGlnThr---AsnPheGlnAspLeuLysAlaGlnHisAspSer---Leu 1266
 Db 3793 AGAGAGATCTTGAACACAGCAATTTGCAAGACATGCACTCAGCAAGAAATTAAGTGCCCTT 3852
 QY 1267 Lys-----GlnAspLeuSerGln 1272
 Db 3853 AAGAGCTGGAATATAGATCGGAGAGAAAGTATTTTCAGGCGCTCATGAGTTGCAACA 3912
 QY 1273 AsnIleGlnGlnSerIleGlnThrGlnAspGlnLeuArgAlaGlnGlnGlnLeuArg 1292
 Db 3913 AGTCAAAAGCAGCAATCACACCTTCACTGCTCTCCTCAAAACAAATGAAAGAGTGAAT 3972
 QY 1293 GlnGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuAspLysSerValGlyIle 1312
 Db 3973 GACCTGAGAGAAATATGTAATATGTAATGCTGAGGCTGAAAGATGATGCTGTAACCTGAGCTG 4032
 QY 1313 SerSerProAsnHisAspAlaValAlaAsnGlnLysValSer----- 1327
 Db 4033 AATGATTCAGAGTCAAGAT 4092
 QY 1328 ---LeuGlnGlnValAsnSerLeu-----GlnSerGlnMetLeuArgGlnGlnValAsp 1344
 Db 4093 CTACTAATATGAAATTAAT 4152
 QY 1345 Gln----- 1345
 Db 4153 GAAGCATATACAGAGAGTGAATTTGTTGAACAACCAATGACAGACACCTGTGTCTTG 4212
 QY 1345 ----- 1345
 Db 4213 GCTCCATTGAGAGAGATTAATCTCTACAGACACTTGACATTTGCAGACAAAGATTCAA 4272
 QY 1346 -----LeuGlnThrSerCysLysAlaLeu 1353
 Db 4273 ATGCATTTCCCGAATTTGCAAGAGAAATTTCTTAATCTTAAGTGAACACAAATTTTA 4332
 QY 1354 -----ValSerGlnLeuGlnLeuLeuArgAlaHisValLysSer 1366
 Db 4333 CATGATCAGACACTGATGAGAGCTCTAATAATGCTGAGAGCTGAGACCTATGTGACTCA 4392
 QY 1367 ValGlnGlnLysLeuGlnIleThrLysLysLeuAsnGlnLysGlnGlnIleLeu 1386
 Db 4393 TTTAAGGCCGAAATTTGCTTGTCAACGAATCTGGAACATTTTCAAGGTGACTGGTG 4452
 QY 1387 GlyLysSerGln-----GlnSerGlnValLeuLysSerMetLeuGlnAsnLeuLys 1403
 Db 4453 AAGGAGATGAGCTGCGCTTGAGAGAGGAGGCTGTCATCTCCTGATCTCTGTGTG 4512
 QY 1404 GlnAspAsnAsnLysLeuLysGlnGlnAlaGln----- 1414
 Db 4513 CCTACAGACTCTAGCTTTACAGAGTTTGGAGAGCTCCCTTTTACAGAGCTCTTTAGAA 4572
 QY 1415 -----GlnThrSerLysGlnAsnGlnPhe 1423
 Db 4573 CAGACAGAGATATGCTCTTTTGAATTAATTAAGAGGCTGTTTCAGCAACCACTGTC 4632

QY 1424 SerLeuGluValPhe----- 1429
DB 4633 AGTGTAGTGAATTTTGGACGAGCTGCAGAGAGAAATCTGACCAGAAAGAAC 4692
QY 1430 ---SerGlySerGlnLysLeuValAspGluLeu-----GluValLeuLysAla 1444
DB 4693 CCTTCGGCCCCAGGAGGGGTGTTGAAGACTTGAGTCCCTCTGTGAGAGTGTACCGGAG 4752
QY 1445 GlnLeuLysAlaAlaGluGluArgLeuLys-----LysLysAspArgPyr 1460
DB 4753 TCCCTCGAGAGCTAGAGAGAGAAATGAAAGTCAAGGATTAATAAATAAGAAAT 4812
QY 1461 PheGluLeuValGlnThrAlaSerThrAsnLeuValGlu---GlyLysLeuGluThrPro 1479
DB 4813 CAAGAGCTCGACAGATTATTAAGTTCTGAAGGCAAGAGCTTGACTGCTTGAAGAGAG 4872
QY 1480 LeuGlnAlaAspHisGlu-----GluAspSerIleAspArgSerGluGluMetGlu 1497
DB 4873 TATTGTGAGAAATGAAACACTGGCAACAGAAAGCTGACAGCGTACTCTGAGATGGAG 4932
QY 1498 IleLysValLeuGlyLysLeuGluArgAsnGlnTyrLeuLeuGlu----- 1513
DB 4933 TCCAGATTGGCGGAGAAAGAAACAGACGAACTGTCACTGAGCTGGAAGTACGA 4992
QY 1514 ArgLeuGlnGluGluLysLeuGluLeuSerAsnLysLeuGluLysLeuGluMet 1533
DB 4993 CGACTCCGCTACAGAGCTGTGACTTAAGTCTCGG----- 5028
QY 1534 GluThrSerValLeuLeuLysAspAspLeuGlnGlnLysLeuGluSerLeuLeuSerGlu 1553
DB 5029 -----TCTTGGCTTGGCATGACAGACAGAAAGATGCTATTTCAAGGC----- 5067
QY 1554 AsnIleIleLeuLysGluAsnIleAspThrThrLeuLysHisLysSerAspThrGlnAla 1573
DB 5068 -----CGAAATGAGAGCTGTGACATCAAAAGAACATACTCTGAGAACTACAGAA 5118
QY 1574 GlnLeuGlnLys-----ThrGlnGlnLeuGln 1583
DB 5119 AGAACACCAAGCATGATGTCATCAGATTGTGCTAAAGATGCTCAGACAGACCTCAT 5178
QY 1584 Leu-----AlaLysAsnLeuAlaIleAlaAspAspAspCysProIle 1598
DB 5179 CTAGACATTGAGAAATTAATCTGAGACTGTGCATTGAAACCAAGAGAGAGCTCTGG 5238
QY 1599 ThrGlnGlnLysGluThr-----SerIle 1606
DB 5239 GAAACAGTCCCGAGATACCAATTATGAGCTCCAGGGAGAAATAAACCAAGGCTCTCA 5298
QY 1607 AspCysValHisProLeu----- 1612
DB 5299 GAATGCAATTTGCAATTTCTATTCTGTGCTTAATGCTTTGGTACCTATGATTTCTGTG 5358
QY 1613 ---GlnGluLysIleLeuLeuLeuThrGlnGluLeuHisGlnLysThrAsnGluGln 1630
DB 5359 GGAATACAGAGAAATATCTCAATATCTCACTCGGGTAAAGACATCAATAGAGAT 5418
QY 1631 GluLysLeuLeuHisGluLysAsnGluLeuGlnAlaGlnValGluLeuLysCysGlu 1650
DB 5419 TTGAGATTACTTCATGTGATAGAGACCGTAGACA-----AAA 5457
QY 1651 ValGlnHisLeuMetLysSerMetIleGluSerLysSerSerLeuGluSerLeuGlnHis 1670
DB 5458 GTTGAAGTTTGGTAAATGAATGAAGAAATTAAGACTCAAAACTC----- 5502
QY 1671 GluLysHisAspThrGlnGlnGlnLeuLeuAla-----LeuLys 1683
DB 5503 ---CATTTACAGAGAGTACACTAATATGACCAAAATTTGAAGCATGATGAATGGAA 5556
QY 1684 GlnGlnMetGlnValValThrGlnGlnLysGluLeuGlnGlnThrHisGlnHisLeu 1703
DB 5557 AAAATAGTTGGGAACTTAAGAAAGAAACTCAGATTTAAGTGAAATAATGGAAATATTTT 5616

QY 1704 ThrAlaGluValAspHisLeuLysGluAsnIleGluLeu-----GlyLeuAsnPheLys 1721
DB 5617 TCTGTGTACACCAGAGATTACTCCAGAGAGTAGAAACTTCTGAAGGCTCAATTTGTAT 5676
QY 1722 AsnGluAlaGln---GlnLysThrThrLysGlnGlnCysLysLeuAsnGlnLysLys--- 1739
DB 5677 TTAGAAATGACATGACATTAATCATCAGCTGAAGATATTGGAGATATATGTGCCAAAGTGTG 5736
QY 1740 -----GluLeuGlnGlnSerGlnHisArgLeuGlnCys 1750
DB 5737 AATGACAGCTGGAAGAGAGATTTTGTGATGTGGAATAATGAGCTGATGATGATGATCG 5796
QY 1751 Glu-----IleGlnGluLeuMetLysSerLeuLys 1760
DB 5797 GAGAAAGCTAGCATTTGAGCATGAGACCCCTTACCTGAGGCTGACTTGAAGGTACTTCA 5856
QY 1761 AspLysGluSerAlaLeuGlnThrLeuLysGluSerGluGlnLysValIle---AsnLeu 1779
DB 5857 ACAGAGAAAGCTATTGTGAAATAAGACAAATGAATAATGAAGAGATTTATGTGCGCTT 5916
QY 1780 AsnGlnGluMetGluMetValMetLeuGluMetGluLysLeuLysAsnSerGlnArgThr 1799
DB 5917 GAAAGAAAGCTCTCAGTGTGACAAAGTGAGAAACAGCTTCTGAGAAATTAATGATACT 5976
QY 1800 ValIleAlaGluArg-----AspGlnLeuGlnAspArgLeuArgGluSerValGlu 1816
DB 5977 ATGTCAAAAAAACCAAGCGACCTGATCATGTTGTGTGAAAAATGAAGACAAACACAA 6036
QY 1817 MetSerIleGluThr-----GlnAspAsp 1824
DB 6037 ---GAGCTTGAATCTCATCAAAAGTAGTGTCTCATTTGCTATTCAGGTGGACAGGACAG 6093
QY 1835 LeuArgLysAlaGlnGlnAlaLeuGlnGlnLysAspLysValGlnGluLeuThrSer 1844
DB 6094 GTGAAGAAAGACGAGACTCTCTCAGACTTTGTCTGTGATGATGATGATGATTAATA 6153
QY 1845 GlnIleSerValLeuGlnGlnLysIleSerLeuLeuGlnLysGlnMetLeuThrVal 1864
DB 6154 GACAAAACTCATCTCCAGAGAAAGCTCCAGAGATTGGAAAGAGACTCAGAGCACTGTCT 6213
QY 1865 AlaThrValLysGluThrLeuSerGluArgAspLeuAsnGlnSerLysGlnHisLeu 1884
DB 6214 TTGCAAAATGTGAGCTGGAACCAAAATTCACACATCAATTAAGAAAGAAATTCCTT 6273
QY 1885 PheSerGluIleGluThrLeuSerLeuSerLeuLysGluLysGluPhe----- 1900
DB 6274 GTCAAGAAATGTGAAGCTGTGAGGCCAGACTGAGTGAATCAGATTATGAAGAAAGCTGAT 6333
QY 1901 -----AlaLeuGlnGlnAlaGluLysAspLysAlaAspAlaAlaArgLysThrIle 1917
DB 6334 GTCTCCAAAGCCTTGGAGCCGCACTGTGGTGAAGAAAGTGAATGATGAGCTGAGC 6393
QY 1918 AspIleThrGluLysIleSerAsnIle----- 1926
DB 6394 TCACACAGAGAGAGTCACTCAGCTGAGAAAGCATGAGAACTGAGAACTGAGATTCGCAAT 6453
QY 1927 -----GlnGlnGlnLeuGlnGlnAlaThrAsnLeuLysGluThrLeuLysGlu 1943
DB 6454 GAGCCCGATGAAGAAAGACAGCTGACATCGAGAGAACTGAAGAAAGCCGAGCGGAG 6513
QY 1944 ArgGluSerLeuIleGlnCysLysGlnGlnLeuAlaLeuAsnThrGlnHisLeuArgGlu 1963
DB 6514 AATGATTCACAT-----AAGGATTAAGTGTGAAGACCTTGAAGGAATTTGACAGATG 6564
QY 1964 ThrLeuLysSerLysAspLeuAlaLeuGlnLysMetGlnGlnLysArgAspGlnAlaAla 1983
DB 6565 TCAGAGAAAGAAACAGAGACTGATGAT-----CTTGATGCCGAGAAATTCACAAACA--- 6615
QY 1984 AsnLysValIleAlaLeuThrGlnLysMetSerSerLeuGlnGlnLysAsn 2003
DB 6616 -----GAGATGAGACTCTTAATAACCAAAATTAAGAAAGATG 6651
QY 2004 ValThrThrLeuLysGluGlu-----GlyLysLysGluThr 2016

```

Db 6652 GCCAGAGCCGTGAAAGTTTGTGAAATTAGACCTGTCACGTTAAAGCTCGAAGAAAT 6711
OY 2017 PhetylLeuGlnArgProSerLysGlnGlnSerSerGlnMetGlnLeuArgGlu 2036
Db 6712 CTGCAAAACAAATACAGAAAAACAGGTCAGTTGTCAGACTGAGCAATGCTCTCT 6771
OY 2037 SerLeuLysThrLysAspLeuGlnLeuGlnLysGlnLysGlnLysSerGlnLysThr 2056
Db 6772 TCATTAAAGTCTGTAGAAAGAAAGAGCAACAGATACAGATCAAGAAAGAAATCT 6831
OY 2057 -----AsnGlnLysAsnLeuThrAlaLysLysSerSer 2068
Db 6832 AAAATGCACTGGAGATGCTTCAGAAATCAGTTAAAGAGCTAAATGAGGAGCGAGCC 6891
OY 2069 Leu-----GlnGlnLysLysLeuGln-----AsnAlaSerLysLeuAsnGlnAlaLysSer 2085
Db 6892 TTGTGTGTGACCAAGAAATTAATACAGCCAGACAGACAGACTAGACCCACCAATAGAG 6951
OY 2086 GluArgGlnAsnLeuArgLysSerLysGlnGlnLeuValSerGlnLeuGln----- 2102
Db 6952 GAAGAGCAGTCAGTAAGAAATAGATGAAAGCTGAGAGCCGCCCTAGAAAGCTGATGAA 7011
OY 2103 -----GlnLeuSerLeuThrLeuLysSerArgAspHisAlaPheAlaGlnSerLysArg 2120
Db 7012 AAGAGAGAGCTGTGTCTTACAAACACTGAG-----GAAAGTGACAT 7056
OY 2121 GlnLysAspGlnAlaValAsnLysLysAlaSerLeuAlaGlnGlnLysLysLysLeuThr 2140
Db 7057 CATGCAGATTACTTAAGGGTAGAGTAGAGACACTTCAAGAGAGCTAGAGATAGCCAGG 7116
OY 2141 LysGlnMetAsp-----GluPheArgAspSerLys-----GluSerLeu 2153
Db 7117 ACAAAACCAAGACATGAGCTCTGAGCGAGAAATTCAAAGAGAGAGTAGAGACCCTA 7176
OY 2154 GlnGlnGlnSerHisLysSerSerGlnLeuLysCysThrLysThrGlnLeuGlnMet 2173
Db 7177 AAGCAAAATATAGAGGATGACCCCAAGTGTGAGAGCTGTGAAATTAAGATGTGTACT 7236
OY 2174 LeuLysGlnGlnLysGlnAspLysAsnLysLeuAlaLysValLysGlnValAsp 2193
Db 7237 ATAAAGTCAGAAAAAAGAAATCTGACAAATCAATTAACAAAAAGCAAGCCAAATCTC 7296
OY 2194 GlnLeuLeuGlnHisLysSerSerLeuLysGlnLeuAspGlnLysGlnMetGlnLeu 2213
Db 7297 GAAATTA-----GAAATA 7308
OY 2214 ArgAsnGlnLysLeuArgAsnLysGlnLeuLysGlnLysMetAspLysMetGlnLysGln 2233
Db 7309 ATAAATTCATCATTTGAAAT----- 7329
OY 2234 LysSerValLeuArgLeuMetGlnAsnGlnProGlnGlnGlnLysAspValAlaGlu 2253
Db 7330 -----ATTTTGCAGAAAAAAGCAAGAGAAAGTACAGATGAAGAAAAATCAGACT 7383
OY 2254 ArgMetAspLysLeuLysSerArgAsnGlnGlnLysGlnLeuMetGlnLysLysSer 2273
Db 7384 GCCATGGAAGTGTTCACAAA-----CAATTAAGAGCTCATGAGAGAGTGCGCA 7434
OY 2274 AlaValLysSerGlnGlnHisThrLeuLysSerSerLeuSerSerGlnLysGln 2293
Db 7435 GCCCTGATATATGACAA----- 7452
OY 2294 ThrGlnAlaHisLysHisCysMetLeuAsnLysGlnLysSerSerThrLeuSer 2313
Db 7453 -----GAAGCCGTGAG-----GCCAAAGAGCAAGATCTTAGT 7485
OY 2314 ArgSerPheLysSerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeu 2333
Db 7486 AGTCAAGTAGAGTGTCTTGAACCTTGAAAGGCTCAGTTGCTACAAAGCCCTGATGAGCC 7545
OY 2334 LeuAsnLysPheLysValValLysArgThrAlaAlaValLysLysLysPheSerLeuLys 2353
Db 2353 LysAspThrAlaValLysArgThrAlaAlaValLysLysLysPheSerLeuLys 2353
Db 7546 AAAAATATATATGTTTGTGCAATCTCAGT-----AATGCCCTCAT 7590
OY 2354 LysAspThrGlnLysAspLeuAlaAlaGlnLysArgHisAspGlu-----Leu 2370
Db 7591 CAAGAGTAGAGTAGGCAAGCAGCAAGCAAGTGGAGAGAGAGTAGAGAAATCAGTAGCTG 7650
OY 2371 ArgLeuLysLeuGlnCysLysLeuGlnHisGlyArgLysThrSerAspSerAlaSerGlu 2390
Db 7651 AAAAATCAAAATTCAGAACCAAGAGCAGCTGTCTTAATGTCCTCCAGGTGGAAGAGAG 7710
OY 2391 GlnLeuLysPhe-----CysGlnLysLeuPheLysLeuLysLeuLysLeuLys 2406
Db 7711 CACCAACTTTGGAGAGAGCAAAACTTAGAACTGAGAAATGTGACAGTGAATTTGAGCAG 7770
OY 2407 LysAlaAsnLysLysLeuGlnSerValGlnAspPheSerGlnValGlnValPheLeuAsn 2426
Db 7771 AAGATCCAACTGCTACATCAAAATGCTCTTGTGAGAGACACATTAGAAAGTGTCCAG 7830
OY 2427 GlnValGlySerThrLeuGlnGlnLysGln-----HisLysLysGlyPheMetGlnThr 2445
Db 7831 AGTTCTTACAAAGATCTAGAGAAATGAGCTTGAATTCACAAAATGACAAAATGTCCTTT 7890
OY 2446 LeuGlnLysPheLys-----AspLeuHisVal 2454
Db 7891 GTTGAAGAAATTAACAAATAGCTGCAGAGAAACTGAGCTGCAGAGGAAATGCAATGAC 7950
OY 2455 AspAlaLysLys-----LeuSerGlnLysMetGlnGlnLysAsnArgGlyAla 2471
Db 7951 ATGGCAGCAAGAAACAGCAGACGCTGCAGAAAGAACTAGTGGAGAGAAATTAAGCTAGCT 8010
OY 2472 SerThrLysLeuLeuThrLysArgLeuLysAlaValAlaGlnSerLysLysGlnArg 2491
Db 8011 GGAGAGTTGACAGTACTGTTGGAAGAAATGAAGAGC-----AGCAAGATGCA 8058
OY 2492 GlnLeuThrValLysLeuAsnGlnPheGlnAlaLysLeuGlnLysLysGlnLysAsn 2511
Db 8059 -----TTGAAGAGCTCACATGAAAT 8082
OY 2512 LysGlnLeuMetArgArgMetGlnHisGlyProSerLysSerValMetGlnGlnGlu 2531
Db 8083 AGTGAATTAAGAAAG-----AGCTATATGTCATGCAACAAAGAC 8121
OY 2532 AsnAlaArgLeuLeuGlyLysLysThrValGlnAspLysSerLysLysLeuGlnSer 2551
Db 8122 CAGCTGGAAGAAAGAGG-----AAGTGAGAGGAAATAGCTGATATACAGCTA 8172
OY 2552 ArgLysMetLeuGlnAsnGlnLysLeuValLysAspAlaMetHisLysGly 2571
Db 8173 CGG-----CTTCATGAAGCT 8187
OY 2572 GlnLys-----ValAlaLysLeuGlnAspLysLeuSerArgAsnAlaGlnLys 2589
Db 8188 GAAAGAGAACACAGCAGCTTGTGCTTTGGAC-----ACAAACAAAGATGAAGTAGAA 8241
OY 2590 LeuAsnAlaMetGlnValLysLeuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLys 2609
Db 8242 ATCCAGACATACCGAAGAAATGACTTTAAAGAAAGATGCTCAGTTCACAGAAAGCTG 8301
OY 2610 GlnLysGlnLysLeuGlnLysMetValAlaLysGlyAlaValProLysLysGlnLys 2629
Db 8302 GAGATGACCTTTAAAGTCTAGT-----AAGAGAGCTC 8337
OY 2630 AspAsn-----LeuLysThrLysValValLysLysLeuMetGln 2642
Db 8338 AATTAATTCATTAAGAACTACTACTAGATTTTGAAGAAATGAAGAAACAGATGAGC 8397
OY 2643 LysLysLysThr-----SerLysAlaThrAspGlnGlnLys 2654
Db 8398 AATCTAAATATATTAATCAGTTGAAGAAAGAAATGACGTCACAGGAGAAATGAAG 8457
OY 2655 AlaLysLeuLysSerLys-----LeuGlnAspLysGlnLysLysLysArgArgLeuLys 2672
Db 8458 TTGTGATCAAAATCTGTAAACAGCTGGAAGAGGAAAGAGAGATTAAGTCAAGAAAGACTC 8517

```

```

GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Glatz, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoersht, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERA
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 12313
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 12031, 12102, 12264
OTHER INFORMATION: n = A,T,C or G
US-10-171-311-7

Alignment Scores:
Pred. No.: 2,34e-56 Length: 12313
Score: 1024.50 Matches: 716
Percent Similarity: 37.198 Conservative: 648
Best Local Similarity: 19.528 Mismatches: 1177
Query Match: 6.948 Indels: 1127
DB: 9 Gaps: 153

US-09-150-867-1 (1-2954) x US-10-171-311-7 (1-12313)
QY 70 ArgSerAlaLeuGlnGlyTyrAsnGlyThrIlePheAlaTyrGlyGlnThrSerSergly 89
||||| :|
51 GSATCCGCCAGTAG-----CGCGAGACGTGCTTCACACTTGGCGGGGGAGCCCGCA 10
90 LysThrTyrThrMetMetGlyThrProAsnSerLeugly-----IleIleProGln 10
105 -----CCGATCGCGCTCTAGGCCCTGAGAGCTTGAGCTGCCCTCC 14
107 AlaIleGlnGluValPheLysIleIleGlnGluIleProAsnArgGluPheLeuLeuArg 12
141 CACCTCCGT-----CCAAATCGACTTTCCTTCTATC 17
127 ValSerTyr-----MetGluIleTyr-AsnGluThrValLysAspLeuLeuCyAspAs 14
174 CCCAACCAACCCCTCAACCCCTGTTTCCCGCGCTTCTTGAGAGGCGCAATGAGAGACA 23
144 PArgArgLysLysProLeuGluIleArgGlu-----AspPheAsnArgAsnVa 16
234 GGAGAGACAGAAAGAGCTGGAGGCGCGCAAGGCCAAGCTTCCCGAGTTCCACAAGAA 29
160 lTyrAlaAlaAspLeuThrGlnGluLeuValMetAlaProGlnHisValIleGlnTrpI 18
294 AGCTCAGTCGATGGGCGAGACTCT-----TC 32
180 eluLysGlyGlyLysAsnArgHisTyrGlyLysThrLysMetAsnArgHisSerSerAr 20
321 CAAAGAACGAAAAAAGAG-----AAAACGTCAAGCAGTAAACATGATGTCTC 37
200 gSerHisThrIlePheArgMetIleValGluSerArgAspArgAsnAspProThrAsnSe 22
:::|::|

```

Db 372 AGCAGCAGCATGATTGATATGATCATCATCAGTGTATGCAATATGATTAATAGTTC 431
 QY 220 rgluasncysapgllyalvalmetvalserhisleuasnleuvalaspheulaglyse 240
 Db 432 TCAGAGAGTAGAATCAACTGTGATTCCTGATCAATATGAGAACTGTACATAGTGG 491
 QY 240 rgluaralaserlgnthrlyalagllyvalargyleus----- 254
 Db 492 AGAATAACACAGTCATGAGCAGGCTTCTGTGCAACTGGAAGGAAATTTCAACCAC 551
 QY 255 -----gluclycysasn-----lleasnargserleupheilleuglyglvalillely 271
 Db 552 ACCAGATGACCTGACATGAGTAATGTTGTCAGCTTTGTGATG----- 597
 QY 271 slysleuseraspgllyglinalaglygphelieasnlyrargaspserlyseuthrar 291
 Db 598 -----AGAACAGGAAGCCCTCAAA 617
 QY 291 glieleuglnaanserleuglyglasnalalyslthvalillecystrhlerhpr 311
 Db 618 TTTATTAAAGGAAGAATTTGGT----- 642
 QY 311 ovalserpheaspglunthrleuserthrleuglnphealaserthralalyslthvalar 331
 Db 643 -----GTTGATGATCTTATTC-----GAAACAGAGGACACA 674
 QY 331 gasnthrprohisvalasncluvallleuaspaspglulaleuleulysarglyrargly 351
 Db 675 AGACAGCCGACATCTGAGATGATGATGAAAGTGGATGGCTGGGAAAGCAGCAT----- 729
 QY 351 sglulieleuaspheulyslysglnleuglnleuasnleuglnuserserglunthrlyal 371
 Db 730 -GAGATTGAAGAGCTTAACAGAGAGCTGGAAGAAATGAGGTGTACTATGAGCAGAGG 788
 QY 371 a-----glnalmetlalysglngluhls-----Thgl 381
 Db 789 ACTGCAGCAGTTACAGAGATTGAGCTGCCATTAACAAAGAGATGATTAATTAACCA 848
 QY 381 nleuenuagluillelysglnleuhslysgluarvglunasparglilethrhisleuth 401
 Db 849 GCTCAGCTGATATTACAAACAGAGAGAGAAAGAGAGAGAGAGAGAGAGATTTT 908
 QY 401 rsnillevalvaliaserlgnleuglnleuaspasnlyrvallyslarglysar 421
 Db 909 AGAGTTG-----ACAGAACAGGTCAAAATTAACGATTCATTT 947
 QY 421 gatgvalthrtrpalaatproglyllyleuglnanserleuhsilaserglyalseras 441
 Db 948 TCAGCAATTACAGCTAGTAAGAACTGTGAAACAGCAGCTCATAGTACAGAGCTGAGA 1007
 QY 441 p-----pheaspmetleuserar 447
 Db 1008 CTTACTACAGCAACACAGATCTCTACATCAACAGCAGCTTGAAGAACAGACA 1067
 QY 447 gleuproglyasnpheserlys----- 454
 Db 1068 CTTATTGAGATATATCAGAAAAGAAAGAGACCTTCACAAATGCAATTTAGTTCTTCA 1127
 QY 455 -----Lysalalyspheserlproserpheprogluileaspaservalcysth 473
 Db 1128 AGAGAAATTTAAAGTATATGAAATGGAACAAAGATTAATAAAGTGAAGAACTCAATTAAGA 1187
 QY 473 rglupheseraspheaspapalaleusermetleaspsersasglylilasapalagl 493
 Db 1188 AGAATAACAGGAAAGAGACAAATCATGAAATTTAAACACAAATATATAGAACAGA 1247
 QY 493 u-----trpasnleualaserlysvalthrhisarglyulyslth-----serie 508
 Db 1248 AAAAGAACTCTTGAGCTTAAGAGATTAATTAACAACCTGCTGATTAATTTACTAGGAGATT 1307
 QY 508 uhsiglnsermetileaspheglyln-----lleasars 520
 Db 1308 ACAAGAAACAGATTGTGCAAAAGAACCAAGAAATTAATAACATGAATTTAGAGCTGACTAA 1367

QY 520 pservalglnphehisaspserlysgluasnleuglnleuasnlyrleuprolyaspse 540
 Db 1368 TTCTTAAGCAAAAAGAGACAGCTCTCTGAAAGAAATTAACAGCTTAATG----- 1416
 QY 540 rglaspmetltaglucysarglys-----Alaserpheglulysgluul 555
 Db 1417 -GGGACAGTCGGAAGACTTCGAGAGAAATCATTAAGACAGCACTGTGAACTGATAT 1475
 QY 555 ethrser-----leuglnglnleuglnserlys-----gluclyglulysgluleuva 573
 Db 1476 AGTAACAAGATGGAACAGAAACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1535
 QY 573 lginserphegluleuyllealagluleuglnleu-----servallysal 591
 Db 1536 TGAGATGTATGGCAGACACATATGTCAAATGAAACAGAAATTAATTAAGACACAGATGC 1595
 QY 591 alyasnleuglumetvalthraserlth-----gluhslerlleasnagl 608
 Db 1596 ACAGATGAGAAATGAAACAGGCAATAGGAGAGAAATGAGAAATGCTTTAAGCTATA 1655
 QY 608 uvalglthrasyalglulysgluvalalarglyglumetservalleuglyaspse 628
 Db 1656 TTCAAAATTTACATTAATGACATGACATTAAG-----TTATGATATGT 1700
 QY 628 rgllyrasnalaseranserlserleuglnaspservalaspgllyllyargyleuse 648
 Db 1701 GGCATTAATATGACTGAATTAATTAATGCAAGATTAATCTCAAAAGGAAACCTCAA 1760
 QY 648 rserserhisaspglucylsileuglnllyarglysmetleuglnleuyllevalasple 668
 Db 1761 GGAAGACTAGACATTAATTTAGAGAAAGTGTCTCTACAGAGACACTTGAAGACT 1820
 QY 668 ucluglnpheillegluasleuasnlyssergluasnaspglulysserl 688
 Db 1821 TGTGTAAGATTTGAGCTTTTCAAGAGAACATGACAGCTAGACAGCAATTAAG 1880
 QY 688 uglinaspphemet-----gluserilleuglnleucysglulalmetla 703
 Db 1881 ACAGAAAGATTAATTAATGAACATTAAGCTTCTGTGATGATGAAAGT 1940
 QY 703 aglu----- 704
 Db 1941 TGAGATTTGTTCTGATCTGAATCCAGAAAGACTAGAAATTAACATGAAGAGAGT 2000
 QY 705 -----lysalaasnalaleu----- 709
 Db 2001 TACAATTAACAGATTAATTAATGTAATTTAGAAAAAGAAAGATGCTGTGATGACAG 2060
 QY 709 ----- 709
 Db 2061 AATGCTGAATCAACAGAGCTGAATTAAGAGCTGAGACACAGCTTCTATTATGCA 2120
 QY 710 -----gluileuallaleumetlrgaspasppheasp-----asnillel 723
 Db 2121 CGAAGAGAGCTTTCCAAATCTGAAGAGAGATTTAGAAATTAAGATGACATGAAATTAATGA 2180
 QY 723 eleuglnasnglnthrleu-----lysargluillealaspheglulargyse 739
 Db 2181 AAAACTTAAGTAAATTTAGCATTCATTAATTAACACACATGATGATGATTAACAGATGA 2240
 QY 739 rleuylsgluasnleuglnthrhasngluphegluileleuglyllysglulthrlnlysgl 759
 Db 2241 AATGAGCTCAAAAGATGAACACATGACTTT-----GAAAGGACAAATTTGATTAAC 2291
 QY 759 uhsiglnleuglnleuhsilysglulileglyserleuylsleuvalalgluasnagl 779
 Db 2292 TAAGCAATCAATTAATTTGGAATTTCAAGAGCTTAAGATTTTACAGAGCTCT 2346
 QY 779 umetlyrasnglnasleuglnlu-----aspheglulthrlysth 793
 Db 2347 -CTGTAAATTCAAAGTCAAGAAATGACTCTCAAAATCAATGATTAATTAAGAAAT 2405

QY 793 rlyseuleuysglu-----GlnGlu1 801
 Db 2406 TGAATACTCAGACAGAAAGAAAGAGGTACCTGACAGACAGAAAGTCAAGAAATT 2465
 QY 801 egluleu-----AlaGluLeuArgLysArgAlaAspAsnLeuGlnLys 815
 Db 2466 ACAACTTAAACAGAAATGTTAGAAAAACAGATGAAAGAAAAAGAAATGATCTTCAAGA 2525
 QY 815 slyValArgAsnPheAspLeuSerValSer---MetGlyAspSerGluLysLeuGly 834
 Db 2526 AAAATTTCACAACTTGAGACAGAGAAATGACATTTCTTAAAGTGAAGAAAAAGCCCTTGA 2585
 QY 834 ugiuilepheGlnLeuLysGlnSerLeuSerAspAlaGluAlaVal----- 849
 Db 2586 AGACATGTGAAGAAATACATCTCCCTGATGACCAAGAGAAAGAAATGATTTCTTAGACTC 2645
 QY 850 -----ThrArgAspAla-----GlnLysGluCysSerPheLeuArgSerG 863
 Db 2646 CATTAAGTCCAAATCCAAAGACCTGTGTGGAAAAAGAAATGAAATGAAATCTTATAGAGA 2705
 QY 863 uasn-----LeuGlnLeuLysGluLysMetGluAspThrSerAs 876
 Db 2706 AAATGAGACCTCAACACACAACTGATTCAGCTAAATGAAAGATGAAAGAAAGAAAGAA 2765
 QY 876 ntrPyrAsnGlnLysGluLysAlaAlaSerLeuPheGlnLysGlnLysGluThrGluLys 896
 Db 2766 CACTTTTCATTTGCTGAAAAAACTTTGAAAGTAACTATCAAGATTAACAAGAGAGATA 2825
 QY 896 sSerAsnTrpLysLysMetGluAlaAspLeuGln-----LysGluLeuG 911
 Db 2826 TGCCTGCTTCMAAGTAAAGATGATTGAAAGACAGATAAAATTAAGAGAAATTAGA 2885
 QY 911 n-----SerAlaPheAsnGluLeuAsnTrpLeuAsn----- 921
 Db 2886 GTATTAAGTAACTTAAGCTTAATGAAAGAGCTTCATTTGCAAGAAATTAATCCAAAC 2945
 QY 921 ----- 921
 Db 2946 TACAGTGAAGAAAGTCTGCTTGTGATGAAGACAAACTTTGTAGCGAAACATT 3005
 QY 922 -----GlyLeuLeuAlaGlyLysValProArgAspLeu----- 933
 Db 3006 GGAATGGGTGAGGTGTTGAAAGGATCAACAGAACTCATGAAAACTTGAGGTAAAC 3065
 QY 934 -----SerArgValGluLeuGluLysLysValSerGluPheSerGlnLeuGlnLys 951
 Db 3066 CAAGCGAGAGAAATTAAGCTGTGACAGAGACTGTGATCTTTCTGAAACAAATTGAACA 3125
 QY 951 salaleuGlnLysAsnAlaLeuGluAsnGluValThrCysLeuSerGluTrpLys-- 970
 Db 3126 GAACATGGTGAGATTGTTCTTAATGAAGAAAGTTAAATCTTTAAAGCAAGAAAGAA 3185
 QY 971 -----PheLeuProAsnGluValGluCysLeuLysAsn----- 981
 Db 3186 ACAAGTTTCATGAGATGATAGAGAGCTGAAGATCATTTATTAACCAACAGAGGCGAAGAA 3245
 QY 982 -----GlnIleSerLysAlaSerGlnGluLeuMetLeuLeuLysG 995
 Db 3246 TGTCAGTCATGATGATCTACAGTAAGTCTTTTATTAGATGAGAGTGTGACCATGACAAG 3305
 QY 995 nglucly-----GluHisSerAlaSerIleIleSerLys----- 1006
 Db 3306 CAGGGTGCTCGAGAGATCAAGTTCTTAAGTAAATTAAGTTTGGTGAAGATCAAAAT 3365
 QY 1007 -----GlnGluIleIleMetGlnGlnGlnSerGlnGln-- 1017
 Db 3366 AATGCTGAGATAAAGTTCTTTTGAAGAAATATGCTGGAGAGAAATGAAGCAAGA 3425
 QY 1018 -----IleLeuGlnLeuThrAspGluValThrHisThrGlnSer-----LysValG 1033
 Db 3426 ACAGTTGATTTTGGATCCTTAACCATCTGTACAAAGAAATCACTACTAGACAACTCA 3485
 QY 1033 nGlnThrGlnGlnGlnTrpLeuGln-----MetLysLysMetHisAspAs 1048

Db 3486 ACCAAGTGAAGAAATGATTAACCTTCAGAGAAAGACTCAATGATCTTAATTCAGAAAGAAATGA 3545
 QY 1048 pleuPheGluLysTrpIleArgAsnLysSerGluAlaGlu----- 1061
 Db 3546 TTTA-----AGCTACAGATGAGAACCCCAACCATTCCTCTCTCTGCT 3590
 QY 1062 -----AspLeuLeuArgGlu---MetGluAsnLeuLysGlyThr----- 1073
 Db 3591 TTATTCAACTATGTCAGATCAGGTTCTGTAATATATGAAAAATGAAAGATTAAGCTCT 3650
 QY 1074 -----MetGluSerValGluValLysIleAlaAspThr 1084
 Db 3651 TTGCAGTCTTAAGAAAGAGCTTATTTTGTCTCAAGAGAAAGATCAAGAACTTCACAGA 3710
 QY 1084 rlyshLsgluLeuGlu---GluThrIleArg-----AspLysGluGlnLe 1098
 Db 3711 AATACACAGTTAGAACTACAGACATATGMAAACACAGAAACAGCTGATGAAGAAAGCC 3770
 QY 1098 uLeuHis-----GluLysLysTrpPhePh 1106
 Db 3771 TTTACATCTGCTCATTGCAAAACTTCMAAGGACGTCTGAGAAATGTTCTTATTTT 3830
 QY 1106 eGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLys 1126
 Db 3831 ACAGACTTATATCAGTATGTC-----CTTGATGAATTTATATCTCTCTCT 3875
 QY 1126 sLeuValGluGlnLysAsnSerGlnAspProIleGluLeuAsnAspTrp----- 1141
 Db 3876 AAATGTGAAGTAAATGACAGAAAGACAAAGAGAAATTCGTGTATTCATTTCTGAAAAATGA 3935
 QY 1142 -----HisAs 1143
 Db 3936 AGATCCAGAAATTACAGATTTATGATATGAACTTCAGACCTTCAGAAAAATATGACACAC 3995
 QY 1143 nLeuIleAlaLeuAlaTrpGluArgAsnAsnIleMetValCysLeuGluThr----- 1160
 Db 3996 TCTTCTCAACAAAGTAAACAGAAATATACACAAACTCTTGATCTTCAACACGACTAAG 4055
 QY 1161 -----GluArgAsnSerLe 1165
 Db 4056 CAAGATCTGGGCGACAGACAGATGATGAAACTGAAATTTGGAGAAAGAAACCTTCC 4115
 QY 1165 uLysGluGlnVal-----IleAs 1171
 Db 4116 AAAAGAGAAACAGAGTTTATATCAATCCATTTCTAGATGACCAATTGGAAAGACATTGA 4175
 QY 1171 pleuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLysSerAspLeuGlnLys 1191
 Db 4176 TGTCATCATTAAGAAACAGATTATCTTCTGCAACATCTTGAAAAAACTTAACCTGGAAGA 4235
 QY 1191 sProLysGlnAspLeuGlu-----GlnGlyG 1200
 Db 4236 ACAAGTTCAAGAAATTAAGAAAGCTCATATCTCTTTCAGACAAACAGTTGAAAGAAACCTGA 4295
 QY 1200 uValLysLeuLeuLeuGlnMetGluLeuLeuLysGlyHisLeu----- 1214
 Db 4296 ACAAACTATGAGCGCAGAGATCCACTGTTTACAGAAAGAGCTTCAGACTGTTAGTACTC 4355
 QY 1215 -----ThrAspSerGlnLe 1219
 Db 4356 CACGGTTCGGCAACGCTTACCTGTTGATTCGGTGTATTTACAGAAATCTGATGCCACAGAG 4415
 QY 1219 uSerIle----- 1221
 Db 4416 AACATGTACCTCGTAAGCTGTGTGAAAAAGAAATATTGATGTATCAATATAGATTTTCTGG 4475
 QY 1222 -----GluLysLeuGlnLeuGlnLysnLeuGlnGluValThrGluLysLeuG 1236
 Db 4476 TGAATTTGAGTGAAGAGGAAACAAATATATGCTTAAGTTGCTTGAAGAAACAAATACCAAGA 4535
 QY 1236 nThrLeuGlnGlnGlnMetLysAsnIleThrIle-----GluArg 1249

```

Db 4536 ACAATTAGAAAGAGTAGCTAAGTATTGTGTCAATGAGTATAGCATTTGCTCACA 4595
QY 1249 GAAGGluLeu----- 1252
Db 4596 AACGGAACGTCTGATATCTGGGGAAAAGAAATAGTCATCATCAAGACACACA 4655
QY 1253 -----GlnThrAsnPhgGluAspLeuLysAlaGluHisAspSerLe 1266
Db 4656 TGCTGTGTGTACAGCAAGAACATTTATTTATGAAATGAAATTTTCACAGATCAAT 4715
QY 1266 uLysGlnAspLeuSerGluAsnIleGluInserIleGluThrGlnAspGluLeuArgAl 1286
Db 4716 TGGTTTCAGACTTTTGACAGACATGGAT-----GTCAATTTAAAGAAATTTAAAC 4769
QY 1286 aalaglnGluLeuArgGluGln---LysGlnLeuValAspSerPheArgGln----- 1303
Db 4770 ACTTAGTAAAGATTAGAGAACATGGAAGAAATTTTATTCAAATGATGATCCCA 4829
QY 1304 -----GlnLeuLeuAspCysSerValGlyIleSer-----SerProAs 1316
Db 4830 TGATATTACCAATCAAGAACATGTGTGTGACTATTTCAGAAAGAAATGTTCTCCAAAGA 4889
QY 1316 nHisAspAlaValAlaAsnGln-----GluLysValSerIleuGluValAlaAsnE 1333
Db 4890 TAAACATTTATAGTTAGACAGTCTATTTCATGATGAGATTCAGTGTCAAGCATGATGC 4949
QY 1333 r-----LeuGlnSerGluMetLeuArgGlyLysIuArgAspGluLeuGlnTh 1348
Db 4950 TTCTAGACAACTAATCTTCAATGAAAGACAGTTGGAAGATGAGACAGAACTTGACG 5009
QY 1348 rSerCysLysAlaLeuValSerGluLeuGluLeuArg---AlaHisValLysSerVa 1367
Db 5010 ACAATACCAAGAACATCAACAGCAAGCAAGAAATTTGTAAGGACAGCACTATGCGCAAT 5069
QY 1367 lGluGluLysAsnLeuGluIleThrLysLysLeuAsnGlyLeuGluLysGluIleLeuG 1387
Db 5070 GAGG----- 5073
QY 1387 yLysSerGluGluSerGluValLeuLysSerMetLeuGluAsnLeuLysGluAspAsnAs 1407
Db 5074 -AGACAGCGAGAAAGACCAG-----GAAACAGCTTACAAAGAAAGATTTAA 5114
QY 1407 nLysLeuLysGluGlnAlaGluIleGluLysSerSerLysGluAsnGlnPheSerLeuGluG 1427
Db 5115 GAGCTTAATAGACAAATTAGCCCAAGAGATCCCTCATAGTAATGAAGAAACGTGTTTCA 5174
QY 1427 uValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuLysAlaGluLeu 1447
Db 5175 G-----AGAGAGAGGGTCTTTTAGAGAGCTGGAGACCTAAACAGCTGTCTT 5225
QY 1447 sAlaAlaGluGluArgLeu-----GluIleLysAspArgAspTyrPheGluLeuValG 1465
Db 5226 AGCTGGAAGAGAGAGAGCTGTGTGTGTGAGTGCACACAGCAGT-----ACGA 5273
QY 1465 nThrAlaAsnThrAsnLeuValGluGlyLysLeuGluThrProLeuGlnAlaAspHisG 1485
Db 5274 AACACAGAAATGGAATGAAAACCAAGAGAGAGATTGA-----GAACAAACATTTAA 5324
QY 1485 uGluAspSerLeuAspArgArgSerGluGluMetGluIleLysValLeuGlyLysLe 1505
Db 5325 AGAAAAGGAATTTAGCAAAACAACTGAAGATGTGCTCCGAGATTTGTCTAATGA-- 5382
QY 1505 uGluArgAsnGlnIleLeuGluGluArgLeuGlnGluLysLeuGluLeuSerAsn 1525
Db 5383 -----AGCATGTGACTCCAGAAAGCTAAT-----ATATAGACT 5414
QY 1525 sLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLeuLysAsp 1542
Db 5415 TTTGAAGATCTCTTAGAGAGTTGAAGACAAACAGCACTGTAAAGAAACAATGTGTCG 5474
QY 1543 -----LeuGlnGlnLysLeuGlnSerLeuLeuSerGluAsnIleIleLe 1557
Db 5475 CCATGTCCTTGGGATTTCTAGATAGATCTAGATAAAGCCAGTCATCTGCCCAATTTTG 5534
QY 1557 uLysGluAsnIleAspThrThrLeuLys-----HisHisSerAspThrGlnAlaG 1574
Db 5535 GAGGTCAAGAGACAGAGCATCTGTAAGATCATGTGTCCATGAGGAACATCAAGAGTTAC 5594
QY 1574 nLeuGlnLys----- 1577
Db 5595 AGATGAATCCATTCCTCTTATCTGTGAAGTATGATCCCAAGAAATGACATTAAGATGTG 5654
QY 1578 -----ThrGlnGlnLeuGlnIleuAlaLysAsnLeu-----AlaIleAl 1591
Db 5655 GTCAAAAGTAAGTACGAGAGAGAACAGAGCTGTGCACAACGACTGTGAGAGAGGTTCG 5714
QY 1591 aalSerAspAsnGlyProIleThrGlnGlu----- 1601
Db 5715 TGAACCTGAATTAACCCCTGAATTAAGAAACTATATGCTGAACATTAAGCTTCGACTACA 5774
QY 1602 -----LysGlnThrSerAlaAspCysValHisPr 1611
Db 5775 AGCAGCACTGAAAAAATCTCTAGAACCATTAAGCAATTAAGTAAAGTAAAGTCACTGAACATGC 5834
QY 1611 oLeuGluGluLysIleLeuLeuLeuThrGluGluLysGlnLysThrAsnGluGln 1631
Db 5835 GAAAGTGAACAGACAGAGTTGATGCTGATCATTTAGACAGAAACAAAGAACACAGCA 5894
QY 1631 uLysLeu-----LeuHisGluLysAsnGluLeu---G 1641
Db 5895 GTCCCTTAGTCCCAAGAGAACTTCGAGAGCCGCTTATAGAGATCCAGGCCACAGCA 5954
QY 1641 uGlnAlaGlnValGluLeuLysCysGluValGlnHisLeuLysSerMetIleGlu 1661
Db 5955 ACAGCTAGCTGTGAGACTCAGT---AAGGCTGAGGGCTCATGTATGCTGACAGATGA 6011
QY 1661 rLysSerSerLeuGluSerLeuGlnHisGluLysHisAsp-----ThrGluG 1677
Db 6012 AAAAATCTTTTGAAGAGCAAAATTCAGCAAAAATACTATATATGATGCTTTGAGCA 6071
QY 1677 nGlnLeuLeuAlaLeuLysGlnIleMetGlnValValThrGlnGluLysGluLeuG 1697
Db 6072 GGAATGTGTATGTGCAAGTAAACAGGTTCGCAAGAAATGGAGCCAGCAACAGCATCA 6131
QY 1697 nGlnThrHisGlnHisLeuThrAlaGluValAspHisLeuLysGluAsnIleGluG 1717
Db 6132 AGAAGAAAGAAATTAATCTGTCCAGCAAAAAGAAAGAAAGTATGAAA-----GCAGAGCAGCA 6185
QY 1717 yLeuAsnPhyLysAsnGlnAlaGlnGlnLysThrThrLysGlnGlnCysLeuLeuAsnG 1737
Db 6186 C-----CCAGTTGAACAAACATTTACTACAGGA 6212
QY 1737 uAsnLysGluLeuGluInserGlnHisArgLeuGlnCysGluIleGluLeuMet 1757
Db 6213 GACAGAAATAATTAAGAAAGAAACAACTAGAAATGCAAGTCAAGCTGAAAAGAGTACGTCA 6272
QY 1757 sSerLeuLysAspLysGluSerAlaLeuGluThrLeuLysGluSerGluGlnLysVal 1777
Db 6273 TGACCTTCAAAAACAAAGTGAAGCTTAGAA-----AT 6305
QY 1777 eAsnLeuAsnGlnLeuMetGluMetValMetLeuGluMetGluLeuLysAsnSerG 1797
Db 6306 AGATGTGAAGAAACAAAGTCACTAGG---TTTATGAGCTTGAAACAGAAAATTAATCTGA 6362
QY 1797 nArgThrValIleAlaGluArgAspGlnLeuGlnAspLeuArgGluSerValGlu 1817
Db 6363 ACTAATGATTTAAAGCCGCAAAACCAAGCACTTGGAAGAACAGCTTGAAGAAAATTAAGGAAA 6422
QY 1817 tSerIleGluThrGlnAspAspLeuArgLysAlaGlnGlnAlaLeuGlnGlnGlnLysAs 1837
Db 6423 ATTTTATGATGACAGCATTTGACAGACATGAGAGATGATATTCACAG----- 6477
QY 1837 pLysValGlnGluLeuThrSerGlnIleSerValLeu----- 1849
Db 6478 -GAAATACAGAAACATTAAGCAACAGCACTTAAGTGTGTTCTCGATTCCAGCTACATGA 6536

```

QY 1850 -----GlnGluLysIleSerLeuGluAsnGlnMetLeuTyrAsnValAlaThrVa 1867
Db 6537 ACATCAAACTAGAGAGGTTGAACAGTTAGCAAAATCATCTGAAGAAAAAGACAAATG 6596
QY 1867 llyGluThrLeuSerGluThgAspAspLeuAsnGlnSerIleGlnHisLeuPheSerG 1887
Db 6597 CAGTACCTCTTCTCTTAAGAGACAGCTTCAAGGGATATACAGAAAGAAAGATGAAGA 6656
QY 1887 uilGluThrLeuSerLeuSerLeuLysGluLysGlnPheAlaLeu----- 1902
Db 6657 AATGAGAACTGAGAGTTACAGATGAGAACTGAGACGCGCTTCTTGAGAGGACCG 6716
QY 1903 -----GluGlnAlaGluLysAspLysAlaAspAl 1912
Db 6717 AAACACTTTGGAGCTGAGAAAGCTAAACAGAAATTCCTAGAGATGACAAATTCGACGC 6776
QY 1912 AlaArgLysThrIleAspIleThrGluLys---IleSerAsnIleGluGlnLeuLe 1931
Db 6777 TGAACGAGATGCCATGACAGAAAGAAAGAAAGATTCACAACTTAGAAGAGCAATTAGA 6836
QY 1931 uGlnGlnAlaThrAsnLeuLysGluThrLeuTyrGluArgGluSerLeu---IleGlnCy 1950
Db 6837 ACAGTTTAGAGACAGACACTGGGAAAATTAAGATGAAGAAAGTTCAACAATTCATGCAAT 6896
QY 1950 slyGluGlnLeuAlaLeuAsnThrGluHisLeuArgGluThrLeuLysSerLysAspLe 1970
Db 6897 AGAAATACAGAAAAGAAATCTACTACCGCTTACAGAACTTGACACAGAAACAAAT 6956
QY 1970 uAlaLeuGlyLysMetGlu----- 1976
Db 6957 ATTTAAGATGACATGAGAAAGCTGGGCTTGCATTAAGAAATTCATGCTCATCTAC 7016
QY 1977 -GlnGluArgAspLunAlaAlaAsnLysValIleAlaLeuThrGluLysMetSerLe 1996
Db 7017 TCAGACCAACATGCTGCTATTGGGAAATTTGCTCAAAATATACAGGAAAAAGAGTGA 7076
QY 1996 uGluGluGlnIleAsnGluAsnValThrThrLeuLysGluGluGluGluLysGluTh 2016
Db 7077 AATGACCAATTAATGACAAAGTTAGCAAACTCCAG----- 7113
QY 2016 rPheTyrLeuGlnArgProSerLysGlnGlnSerSerSerGlnMetGluLysLeuArgG 2036
Db 7114 -----CAGCACTTAAATTAACACAGATTAACAAAGGTTATGTGAAGAAAAAAGA 7163
QY 2036 uSerLeuLysThrLysAspLeuGlnLeuGlu----- 2046
Db 7164 ACTGATTAAGGATCTTGAACCCCAATGAGATGTTTGATGATGATGATCAAGATGTGAA 7223
QY 2047 -----GluAlaGluLysGluIleSerGluAlaThrAsnGluIleLysAsnLeuThrAla 2065
Db 7224 GAGAAATTAAGAAAGAAAGAAATGAGAGCTCAATGAGTGATGAAAAAATCAACAGGA 7283
QY 2065 sIleSerSerLeuGluGluGluIleLeuGlnAsnAlaSerIleLeuAsnGluAlaValSe 2085
Db 7284 ATTTGGCAAAATTTGAGCAGAGAGACATCAATGAAATGCTCATTCCTCTCAGAA----- 7335
QY 2085 rGluArgGluAsnLeuArgHisSerLysGlnGlnLeuValSerGlu-----LeuG 2102
Db 7336 -GAAGCAGACAGTTTAAACATCAATGATGTGTTATACCTGAAAAAGCTGGCGTTGA 7394
QY 2102 uGln-----LeuSerLeuThrLeuLysSerAr 2111
Db 7395 ACAGCAAGTGAAGAAACCTTAATGAAGAAATGACCTTATGAAAAATGTACTTAAAGAAAC 7454
QY 2111 gaSPHisAlaPheAlaGln-----SerLysArgGluLysAspG 2124
Db 7455 CAATTTTAAATGATGATGACAAACAGAGAAATTTATCAGCTTAAAGAGAGAACT---GA 7511
QY 2124 uAlaValAsnLysIleLeuSerLeuAlaGluGlu-----IleLysIleLe 2139
Db 7512 AAGGTGTGAAAGATTTAAAGCATACAGAGAAATGATGTTAAAGCTGCTATAGATCATCT 7571
QY 2139 uThrLysGluMetAspLysPhe-----ArgAspSerLysGluSerLeuG 2154

Db 7572 GAGCAAAAGCAAAACCTGTAAGTGAAGTATGCTTACAGAGATGCTCTTAATCCCTGGA 7631
QY 2154 nGluGlnSerSerHisLeuSerGluGlu-----LeuCySThrTyrLys 2168
Db 7632 AAATCAGCATPACTTCAATCTTTTGAAGAAATGCAAGAGTTCATTAATTAATTGGA 7691
QY 2168 sThrGluLeuGlnMetLeuLysGlnGlnLysGluAspIleAsnAsnLysLeuAlaGluLys 2188
Db 7692 AACAAGTGTCTACAACTTGAGAGACGCTTATAGTCCAAAGAGCTTACAACTTACCAGTG 7751
QY 2188 sValLysGluVal-----AspGluLeuLeuGln-- 2197
Db 7752 TTATTAACAAATTAAGACATGCAAGAACAGCCAGTTTGAAACAGAAATGCTTCAAAA 7811
QY 2198 -----HisLeuSerSerLeuLysGluGln-----LeuAspG 2208
Db 7812 GAAGATTTGTAACCTTACAGAAATTAAGTGAAGAAAGAAAGTGGCTGCTCTTGTCTGCA 7871
QY 2208 nIleGlnMetGluLeuArgAsnGlu----- 2216
Db 7872 AATCCAACTTGAGCGATTCAGAGATTCAGAAATTCGTAAGTATTCACAAATTC 7931
QY 2217 -----LysLeuArgAsnTyrGluLeuCy 2224
Db 7932 ATCAGAACTGAAAGAACAAATTCAGAAATTTAAATCAGTAAGAGAAATGAGTTGG 7991
QY 2224 sGluLysMetAspIleMetGluLysGluIleSerValLeuArgLeuMetGlnAsnGluPr 2244
Db 7992 GTCAAGATATACGACATTAACCTTGAGAAATTCACAAATTCAGAAAGCCAGTTGTGAAT 8051
QY 2244 o-----GlnGlnGluGluAspAspValAlaGluArgMetAspIleLe 2258
Db 8052 GCATCTGCTGTTGATTTAGAAAAAGAACAGTACAAATTCGACAAAAA---AATGTTT 8108
QY 2258 uGluSerArgAsnGln-----GluIleGlnGluLeuMetGluLysIleSerAlaValTyr 2276
Db 8109 AGAAAAAGAAAGAAAGCTGCTAGAACCTACAGAAAGCTATTTGGGCAATGAGAAAAAACA 8168
QY 2276 rSerGluGlnHisThrLeuLeuSerSerLeuSerSerLeuGlnLysGluThrGluAl 2296
Db 8169 GAGAGCAAAAGAAAGAAAGAAAGAAAGCCCTCAAGATGTTCAAGACAACTGCA 8228
QY 2296 aHisLysHisCysMetLeuAsnIleLysGluSerLeuSerSerThrLeuSerArgSerPh 2316
Db 8229 GCTATTTAT-----AGCAATGAAGAAAGTGATTTTAAATGAAC 8270
QY 2316 eGlySerLeuGlnThrGluHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLys 2336
Db 8271 CGAGGCTCTTAGAGCTGAATCAGTGCCTACCAAGACAGAACTTGCCAGTTTAAAGAAA 8330
QY 2336 sPheLysValValTyrArgThrAlaAlaValLysGluAspHis-----SerLeuIleLys 2354
Db 8331 GGCCTGAAGAACTTCAAGAAAGAGCTTTTGTAAAGAAACAAATATGATCATCTTCACAGA 8390
QY 2354 sAspTyrGlu-----LysAspLeuAlaAlaGluGlnLys----- 2365
Db 8391 AGACTTAAGCCAAAGTTAGGAGTACCTCGCAGAGGCAAGAAAGAAATGTCCATTTTGA 8450
QY 2366 -ArgHisAspGluLeuArgLeuGln-----LeuGlnCysLeu---GlnGlnHisG 2381
Db 8451 AAAAGAAATGAGACTGAGTGAAGTACAAAGAAAGAAAGCCCTGCATGTTTGAGCCACTTC 8510
QY 2381 YArgLysTrpSerAspSerAlaSerGluGlu-----LeuLysPheCysGluI 2397
Db 8511 TATTAAGCTGATGAAGACATTCCTCCAGACAGATGAGCTGAAGATCAGATACAG 8570
QY 2397 eGluPheLeuAsnGluLeuLeuPheLysLysAlaAsnIleIleGlnSerValGlnAspAs 2417
Db 8571 CAATCAGACTCCACAAATTCCTGTTAAATATGACGAAATTAATTAATTAATTAATTAAT 8630
QY 2417 pPheSer-----GluValGlnValPheLeuAsnGlnValGlySerThrLeuGlnGlu 2435

```

Db 8631 ATGTCCTCAGAGAGAGTACTGAATATATCAGTAC----- 8667
QY 2435 ULEUGLHLSLysLysGlyPheMetGlnTrpLeuGlnGluPheGlyAspLeuHsValas 2455
Db 8668 -----TTACTGAAAAAATTGAGAAATGCAAGATCTACATGCT-- 8706
QY 2455 PALALysLysLeuSerGlnLysMetGlnGlnGlnAsnArgArgIleAlaSerThr----- 2473
Db 8707 -----GCTGAATTTTGGACATGGAATCCAGACATATTTCAAGAACTGAAC 8753
QY 2474 -----IleGlnLeuThrLysArg-----LeuLys 2482
Db 8754 CTTAAGAGGGAACACTATGTCCTGCTCAGTACTGAAAGAGAAATGCTGACTTGAA 8813
QY 2482 SALAValAlaGln-----SerLysIle----- 2489
Db 8814 GCGAGGATACATGCTGTGAGAGTAAGAGGATCTCTCATTTCTGAGCTACACATTC 8873
QY 2490 -----GlnArgLysIleThr----- 2494
Db 8874 TGATGCTTACAGACTAGAGAAATATGCTCAGATTCGTGATCAGACTGGGTCAGG 8933
QY 2495 -ValLysLeuAsnGln-----PheGlnAlaLysLeuGlnGlnLysGlnGlnAs 2511
Db 8934 AATTATCTTACACACAGCTCAGGAGATTGACATGACATCAGAAAGCCGAGAGAGAAAG 8993
QY 2511 n-----LysGlnLeuMetArgArgMetGlnH 2520
Db 8994 TGAAGAGCAACAGATTCTTCCAAAGAAATTAAGGATTTACTGAGACTGTCCATA 9053
QY 2520 SHISGlyProSerAla-----Se 2526
Db 9054 TGAAGGATCAGAGTGTCTCTCACTGAGTCCCTATAGTAGAGAGAGACCATTC 9113
QY 2526 rValMetGln-----GlnGlnAsnAlaArgLeuGlnLysIleLeuLys 2540
Db 9114 TATTCACAGAGCTTTCAGACCTTGCTGAGAGAGAGAAAGCTTACTCATATCAATCTC 9173
QY 2540 strValGlnAspLysSerLysLysLeuGlnSerArgIleLysMetLeuGlnLysGlnLys 2560
Db 9174 ATCTTAAGAGATTATATTCAAAGATGCAA-----CTGCAAGAGAGAGC 9218
QY 2560 uAsnLeuValLysAspAlaMetHis-----LysGlnLysGlnVal 2574
Db 9219 CGAGGTTATGATAGTTCTCAATCTCATGAGAGCTTCTCAGACTGGGAGAGTGAATCTCT 9278
QY 2574 lAlaIleLeuGlnAspLysLeu----- 2582
Db 9279 GCTTGCCCTTCAACAGTTTCTTGAAGAGCGTAGTGTTCCTAGTACAGCATTTCCGAC 9338
QY 2583 -----SerArgAsnAlaGlnAlaGlnLeuAsnAlaMetGlnValLys 2596
Db 9339 GAGAGTCAGACGCTCTAGCTACTACAGATGCAAGTTGTTACTTAACGTGTTGGAACGAG 9398
QY 2596 sLeuThrLysLysGlnAspAsnLeuGlnAlaAlaMetLys----- 2609
Db 9399 AATACAAAGAACAGAGGTGAATATCAAGCAGCTATGGAATGCCCTCCAGAAACAGATAG 9458
QY 2610 -----GlnIleGlnAsnLeuGlnLysMetValAlaLysGlnAlaValPr 2624
Db 9459 AAGAGGTTTCTTATCTGAATTTCAAGCAGCTGATGCAAAATGATAGTGGAAAAATTAC 9518
QY 2624 oTyrlLysGlnGlnLysLeu-----AsnLeuLys 2633
Db 9519 TCTGAAAAGAGAACAGAGAGTGAAGAACCAAGCAAGAACTCTTGGAAATATATATACA 9578
QY 2633 strLys-----ValLysIleGlnMetGlnLysLysLysLysLys 2648
Db 9579 GCGAAGACACTCTCAATGCTGTGAGAGCAAGTGCAGCTCAGAGATGAA-----GACAG 9635
QY 2648 sAlaThrAspGlnGlnLysLeuLysSerCysLysLeuLysLysLysLysLys 2668
Db 9636 AGCAAGCGAA-----CTGCAAGAGACAGCTGAGTTCTGAGAAATGCTGCT 9680

```

```

QY 2668 uArgArgLeuLysGlnGlnLysArgArgAlaGlnAlaAspAsnAspThrThrValCysVal 2688
Db 9681 TCGTGAACCTGAAAGAGAGCTTGCACAACTAAATTGGAACTGGAACAACACTCAAGC 9740
QY 2688 lProLysAspPyrGlnLysAlaSerThrPheProValThrCysGlyGlyLysSerGlyLys 2708
Db 9741 ACAGCATTAACACCTTAAAA----- 9759
QY 2708 eValGlnSerThrAlaMetLeuValLeuGlnSerGlnLysAlaAlaLeuGlnLysGlnLys 2728
Db 9760 -----GAATTGGAGCTTTTACAGTTGGAAGT 9785
QY 2728 uSerHisTyrlLysLysLysTyrlHisLysLeuSerArgThrMetSerSerGlnLysPac 2748
Db 9786 TAAAGATTAAGACAGATGAAGTACATTTGCTTATGACACATTTAGCAAGTGAAGAAAA 9845
QY 2748 g-----LysLysThrLysAlaLysSerAspAlaHisSerS 2760
Db 9846 ATCAAGAGAGCTCCAGTGGCTTGGAGAAAGAGAAAGCCAG----- 9888
QY 2760 rHisThrGlySerSerHisArgLysSerProHisLysThrGlnThrTyrlArgHisGly-- 2779
Db 9889 -----TTGGAGCGCAGTGAAGACGGATTAAGAAAGAACTTGAGAGATCTGAAGTTTCACT 9944
QY 2780 -ProValThrProGlnArgSerGlnMetProSerLeuHisLysLeuGlySerProLysLysSe 2799
Db 9945 TGAGAGTCAAGAAACAAAGCAATCTCAGCTAAATCTACTTTTGAACAACAGAACACT 10004
QY 2799 rGlnSerSerThrLysArgValAlaSerProAsnArgSerGlnLysLysLysLysLys 2819
Db 10005 ACTGAAGCAATCCACCAAAATAAGATCAAGACAGAAATGCTTATGATGCCAGTTTTC 10064
QY 2819 lMetSerProGlyLysThrGlyMetHisLysHisLysLeuSerProSerLysValGlyLe 2839
Db 10065 AGAAGAAACAGCTGCAAACTTACAGCTCAGTACTTCTGATCTGAGAAATTTGGANT 10124
QY 2839 uHisLysLysArgAlaLeuSerProAsnArgSerGlnMetProThrGlnHisValLysLe 2859
Db 10125 TCGGGAATAGTATGATACCTTACATAGTGGAGCCGAGAA----- 10161
QY 2859 rProGlyLysThrGlyLeuHisLysAsnLeuThrGlnSerThrLeuPheAspAsnLeu 2879
Db 10162 -----TTGACCGCACAGCTGCGAGCAGATGATGTGAGACAGTCTGC 10205
QY 2879 eSerProCysLysGlnLys 2885
Db 10206 GCCACCTTGCCCTCAGAGG 10225

RESULT 9
US-10-171-311-3
; Sequence 3, Application US/10171311
; Publication No. US20030087270A1
GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kametkar, Shubhangi
; APPLICANT: Glatf, Karen
; APPLICANT: Ganavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936

```


Oy 628 rgltyrAsAlaSerAsnSerAspLeuGlnAspSerValAspGlyLysArgLeuSe 648
 Db 1701 GGCATAAATGACATGAAATATAAATGCAAGATTAAGTCTCAAAAGCAAAATCTCA 1760
 Oy 648 rSerSerHisAspGlyLysIleGluHisArgLysMetLeuGlnLysIleValAspLe 668
 Db 1761 GGAGAACTAGAGACTAATTTAGAGCAAAAGTGTGCTCTACAGACAGAGCTTGAACCT 1820
 Oy 668 uGluGluPheIleGluAsnLeuAsnLysLysSerGluAsnAspLysGlnLysSerSerG 688
 Db 1821 TGTGGAAGATAGAGCTTTTCCAGGAAAGATTCAGAGAGCTAGAGCAATTAAGCTGA 1880
 Oy 688 uGlnAspPheMet-----GluSerIleGluLysGlnLysIleValMetAl 703
 Db 1881 ACAAGAAAGTAACCTTAATGAGACATTAAGTCCCTTAGACAGTGAAGATTTGAAGC 1940
 Oy 703 aglu----- 704
 Db 1941 TGAGATGTTCTGATCTGAAATCCAGAAAGAACTAGAAATTAACATGAGCAGAACT 2000
 Oy 705 -----LysAlaAsnAlaLeu----- 709
 Db 2001 TACAATATTACAGATTAATTAAGATTTGAAATTAAGAAAGAAATGCTGTGTTAGACAG 2060
 Oy 709 ----- 709
 Db 2061 AATGCTGAATCAACAAGAACTGAATTTAGAGAGCTGAGAACAGCTTCTATTAGTCA 2120
 Oy 710 ----GluGluLeuAlaLeuMetArgAspAsnPheAsp-----AsnIleI 723
 Db 2121 CGAAGAGAGCTTTCACAACTGAGAGAGATTTAGAAATTTGACATCGAAATTAATTTGA 2180
 Oy 723 eleuGluAsnGluThrLeu-----LysArgGluIleAlaAspLeuGluArgSe 739
 Db 2181 AAAACTTAAGATTAATTTAGGATTCACATTAACAGCAGATGATGTTTACAGAAATGA 2240
 Oy 739 rLeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGluLysGluThrGlnLysG 759
 Db 2241 AATGACTCAAAAGATGAGAAACCATGCACTTT-----GAAAAGCAATTTGATTAAC 2291
 Oy 759 uHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsnAlaG 779
 Db 2292 TAAGCAGAACTCAATTAATTTTGAATTTCAAGCTAAAGATTTTACAGACAGTCTC 2346
 Oy 779 uMetTyraSngLysAsnLeuGlu-----AspLeuGluThrLysThr 793
 Db 2347 -CTTGTAATTTCAAAAGTCAAGAAATGACTCTTCAATCATGAACTTCAAAAAGAAAT 2405
 Oy 793 rLysLeuLeuLysGlu-----GlnGluIle 801
 Db 2406 TGAATATCTCAGACACAGAAAGAAAGAAAGGCTGACACTTGAACAGAACTTCAGAAAT 2465
 Oy 801 eGlnLeu-----AlaGluLeuArgLysArgAlaAspAsnLeuGln 815
 Db 2466 ACACTTTAAACAGAAATTTGTAGAAAAACAGATGAAAGAAAGAGATGATTTTCAAGA 2525
 Oy 815 sLysValArgAsnPheAspLeuSerValSer--MetGlyAspSerGluLysLeuGlySer 834
 Db 2526 AAAATTGACAACTGAGAGAGAAATGATGATTTTAAAGATGAAAGAAAGAAAGCTTGA 2585
 Oy 834 uGluIlePheGlnLeuLysGlnSerLeuSerAspAlaGluAlaVal----- 849
 Db 2586 AGACATGTTGAAATATACATCTCTGTAGCCAAAGAAAGAAAGATGATTTTCTTAGACTC 2645
 Oy 850 -----ThrArgAspAla-----GlnLysGluCysSerPheLeuArgSerG 863
 Db 2646 CATTAAGTCCAAATCCAAAGACTCTGTGTGGAAAAAGAAATGAATTAATCTTATAGAGA 2705
 Oy 863 uAsn-----LeuGluLeuLysGlnLysMetGluAspThrSerAs 876
 Db 2706 AAATGAGAGCTCAACAACAATGATTAATCAGCTTAATGAAAGATTTGAAAGCAAAAGAA 2765

Oy 876 nTPeTyraSngLysGluLysAlaLaserLeuPheGlnLysGlnLeuGluThrGlnL 896
 Db 2766 CACTTTTTCATTTGCTGAAAAAACTTTGAACTTAATTAAGGTTTCAAGAGTTTCAAGAGAGA 2825
 Oy 896 sSerAsnTyraLysLysMetGluAlaAspLeuGln-----LysGluLeuG 911
 Db 2826 TGTGTCCTTCTCAAACTTAAGATGATTTAGAAAGCAGTAAAGAAATTAACAGAAATTA 2885
 Oy 911 n-----SerAlaPheAsnGluIleAsnTyraLysAsn----- 921
 Db 2886 GTATTAAGTAAGTAACTTAAGACCTTAATGAAGAGCTTCAATTTGCAAAAGAAATTAATCAAC 2945
 Oy 921 ----- 921
 Db 2946 TACAGTAAATGAAAGTCTGCTTTGATGAGACAAACAACTTTGTAGCAGAAACAT 3005
 Oy 922 -----GlyLeuLeuAlaGlyLysValProArgAspLeu----- 933
 Db 3006 GGAATGGGTGGGTGTTGAAAGAGATACACAGAACTCATGAGAAACTTGAGGTAAAC 3065
 Oy 934 -----SerArgValGluLeuGluLysLysValSerGluPheSerLysGlnLeuGlu 951
 Db 3066 CAAGCAGAGAAATTAAGAGCTGTACAGAGAGCTGTGATCTTCTGACAAATTTGAACA 3125
 Oy 951 sAlaLeuGluGluLysAsnAlaLeuGluAsnGluValThrCysLeuSerGluTyraLys 970
 Db 3126 GAACATGCTGAGATTAATGTTCTTAATGAAGATTAATCTTTAAAGCAAGAAAGA 3185
 Oy 971 -----PheLeuProAsnGluValGluCysLeuLysAsn----- 981
 Db 3186 ACAAGTTTATGAGATGATGAGCTAGAGCTGAAGTAACTATTATTAACCAACAGGCGAGAAA 3245
 Oy 982 -----GlnIleSerLysAlaSerGluGluIleMetLeuLeuLysG 995
 Db 3246 TGTACAGTATGATACTCAAGTAACTGTTTATTAGATGAGAGTTGTACACTGACAG 3305
 Oy 995 nGluGly-----GluHisSerAlaSerIleLeuSerLys----- 1006
 Db 3306 CAGGGTGTCAAGAGATCACTTCTTAAGTAAATAAAGTTTGTGGAATCAAAAT 3365
 Oy 1007 -----GlnGluIleMetGlnGluGlnSerGlnGln 1017
 Db 3366 AATGCTGAAGATTAAGTCTTCTTGAATAATGAGCTGTTGAGAGAAAGATTAACCAAGA 3425
 Oy 1018 -----IleLeuGlnLeuThrAspGluValThrHisThrGlnSer-----LysValG 1033
 Db 3426 ACACTTATTTTGGATCATTACCATCTGTAAACAAAGATCATCATTAAGCAACTCA 3485
 Oy 1033 nGlnThrGluGluGlnLysLeu-----MetLysLysMetHisAspAs 1048
 Db 3486 ACCAAGTGAATAATGATTAATCTGAGAAAGAACTCAATGATTAATTAATCAAGAAATGA 3545
 Oy 1048 pLeuPheGluLysTyraIleArgAsnLysSerGluAlaGlu----- 1061
 Db 3546 TTTA-----AGCTACAGATGAGAAAGCCAAAGCATTTGCCCTCTCTGCT 3590
 Oy 1062 -----AspLeuLeuArgGlu-----MetGluAsnLeuLysGlyThr----- 1073
 Db 3591 TTATTCACTCATGTGATCAGTTCTGTAATATGAAAAAGAAAGATTAAGCTCT 3650
 Oy 1074 -----MetGluSerValGluValLysIleAlaAspThr 1084
 Db 3651 TTGCAGTCTTAAGAGAGCTTATTTTCTCAAGAGAAAGAAAGATCAAGAGACTTCAGAA 3710
 Oy 1084 rLysHisGluLeuGlu-----GluThrIleArg-----AspLysGluGlnLe 1098
 Db 3711 AATACACAGATTAGACATACAGACTATGAAACAAAGAAAGAGTATAGAGAAAGCC 3770
 Oy 1098 uLeuHis-----GluLysLysTyraPheP 1106
 Db 3771 TTTACATCTGCTCATTTGAAACAACTTCAAGAGAGCTGTCTGAAGAAAGTTTATTTT 3830
 Oy 1106 eGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProPheLys 1126

```

Db 8831 ACAGACTTATGACAGTGC-----CTTGGATATATATACCTCCCTTT 3875
Qy 1126 sLeuValGluGlyAsnSerGlnAspProIleGluIleAsnAspTy- 1141
Db 3876 AAAATGTGAAGTAATGCGAAGAACAAAGAAATTCGTGTATTCATTCTGMAATGA 3935
Qy 1142 -----HisAs 1143
Db 3936 ACATCCAGATTAACAAGATTATGATAGATTCACAGCTTCAAGAAAATATGACAC 3995
Qy 1143 nLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr 1160
Db 3996 TCTTCTCAACAAAGTAAGAGAAATATACAAACACTCTTGACTCTCAACAGCACTAG 4055
Qy 1161 -----GluArgAsnSerIle 1165
Db 4056 CAAGATCTGGGAGACAGACAGATGCTATGAAACTTGAATTTGAGAAAGAAACCTTCC 4115
Qy 1165 uLysGluGluVal-----IleAs 1171
Db 4116 AAAAGGGAACAGAGATTTTATCAATCCATCTCAGATGACCAATTTGGAAAGACATTGA 4175
Qy 1171 pLeuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluYsSerAspLeuGlnIly 1191
Db 4176 TGTCAATCATATAAGCAAGATTATCTCTCTGCAAGATCTTGAATAAACTAACTTGAGA 4235
Qy 1191 sProLysGlnAspLeuGlu-----GluGlyGlu 1200
Db 4236 ACAAGTTCAAGAAATTGAAAGGCTCATATCTCTTGACAGCAAGCTTGAAGAAACTGA 4295
Qy 1200 uValLysLeuLeuLeuGluMetGluLeuLeuYsGlyHisLeu----- 1214
Db 4296 ACAAAACATATGAGCGAGAGATCCACTGTTTACAGAAAGAGCTTCAAGCTGTAGTAGTC 4355
Qy 1215 -----ThrAspSerGlnIle 1219
Db 4356 CACGGTCCGCCAGACTTACTCTGTGATCGGTGATATACAGATCTGATGACACAGAG 4415
Qy 1219 uSerIle----- 1221
Db 4416 AACAAATGATACCTGGAAGTGTGTGTAATAAAGATATGATGATACAAATAGAGTTTCTGG 4475
Qy 1222 -----GluYsLeuGlnLeuGlnLeuAsnLeuGluValAlaThrGluYsLeuGlu 1236
Db 4476 TGAATTTGAGAGTAAGAGAGAAACAATATCTTAAGTTGCTTGAATAACAATACCAAGA 4535
Qy 1236 nThrLeuGlnGluGluMetLysAsnIleThrIle-----GluArg 1249
Db 4536 ACAATTAAGAAAGAAAGTAGTAGTATGTGTCAATGAGTATAGCATTTGTCACACA 4595
Qy 1249 gAsnGluLeu----- 1252
Db 4596 AACTGAAGTCTAGATATCTGGGGGAAAGAAATACTGCATCATCAAGCAAGACACA 4655
Qy 1253 -----GlnThrAspPheGluAspLeuLysAlaGluHisAspSerIle 1266
Db 4656 TGCTGTGTCTGCAAGAAACAACATATTTATGAAATGAATATATATCAGAGATCAAT 4715
Qy 1266 uLysGlnAspLeuSerGlnAsnIleGluGlnSerIleGluThrGlnAspGluLeuArgAl 1286
Db 4716 TGGTTTCAGACTTTTGAGACAGCTGAT-----GTGAATTTAAAGAAAGATTTAAACC 4769
Qy 1286 AlaIleGlnGluLeuArgGluGln-----LysGlnLeuValAspSerPheArgGln----- 1303
Db 4770 ACTTAGTAAGAGTATGAGAAACATGAAAGAAATTTATATATCAAAATAGATCCCA 4829
Qy 1304 -----GlnLeuLeuAspCysSerValGlyIleSer-----SerProAs 1316
Db 4830 TGATATACCAAGATCAAGAGAGCTGTGTCTGACTATTTTCAGAAAGAAATGTTTCCAAAG 4889
Qy 1316 HisAspAlaValAlaAsnGln-----GluYsValSerLeuGluValAlaAsnSe 1333

```

```

Db 4890 TAAACATTAATAGTAGACAGTCTATCATGATGATTCAGTCAAGCATGATGATGC 4949
Qy 1333 r-----LeuGlnSerGluMetLeuAlaGlyGluArgAspGluLeuGlnTh 1348
Db 4950 TTCTAGACACTAATGTATTAATGAAGACAGATTGGAAGATATAGACAGCAACTGTAGC 5009
Qy 1348 rSerCysAlaLeuValSerGluLeuLeuLeuAlaArg-----AlaHisValLysSerVa 1367
Db 5010 ACAATTCGAAGAACATCAACAGCAAGCAATTTGTTAAAGCAAGCATATATGCGCAAT 5069
Qy 1367 lGluGluGlnAsnLeuGluIleThrLysLysLeuAsnGlyLeuGluYsGluIleLeuGlu 1387
Db 5070 GGAG----- 5073
Qy 1387 yLysSerGluGluSerGluValLeuLysSerMetLeuGluAsnLeuLysGluAspAsnAs 1407
Db 5074 -AGACAGCAGAGAACACAC-----GAACAGCTTCAAGAAAGAGATTAA 5114
Qy 1407 nLysLeuLysGluGlnAlaGluGluThrSerSerLysGluAsnGlnPheSerLeuGluGlu 1427
Db 5115 GAGACTTAATAGCAATTAAGCCAGAGATCTTCATAGATTAATGAAACCTGTTGACA 5174
Qy 1427 uValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuLysAlaGlnLeuLys 1447
Db 5175 G-----AGAGAGAGGCTGCTTTAGAGAGCTGGAAGACATTAACAGCTGCTTT 5225
Qy 1447 sAlaIleGluGluArgLeu-----GluIleLysAspArgAspTyPheGluLeuValGlu 1465
Db 5226 ACCTGGAAGAGAGAACCTGTGTGTAGCTGCGCAACACCACT-----ACGA 5273
Qy 1465 nThrAlaAsnThrAsnLeuValGluGluLysLeuGluThrProLeuGlnAlaAspHisGlu 1485
Db 5274 AACACAGATGGAATGAAGAAACCAAGAGAGATGGA-----GAACAAACATTTAA 5324
Qy 1485 uGluAspSerIleAspArgSerGluGluMetGluIleLysValLeuGluGluLysLe 1505
Db 5325 AGAAAGGAATTAAGACAGAAACCTGAGATGCTGCTGCTGATTTGCTTAAGCA-- 5382
Qy 1505 uGluArgAsnGlnTyLeuLeuGluArgLeuGlnGluGluLysLeuGluLeuSerAsnLys 1525
Db 5383 -----AGTATGACTCCAGAAACCTAAT-----AATAGACT 5414
Qy 1525 sLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLeuLysAspAsp----- 1542
Db 5415 TTGGAAGATCCCTTGAAGTTGTAAGACAAACAGACAGCTGTGAAGAAACAATTCGTGC 5474
Qy 1543 -----LeuGlnGlnLysLeuGlnSerLeuLeuSerGluAsnIleIleLe 1557
Db 5475 CCATGCTTGGGATTCATGATATAGATCTAGTAAAGCCAGTCATCTGCCACCTAATTTG 5534
Qy 1557 uLysGluAsnIleAspThrThrLeuLys-----HisHisSerAspThrGlnAlaGlu 1574
Db 5535 GAGGTCAGAGACAGAGCATGTGTAAGTCATGTCTGAGAGAACATATCAAGAGTTAC 5594
Qy 1574 nLeuGlnLys----- 1577
Db 5595 ACATGAATCCATTCCTCTTATTCGGAAGTGATATGCCAAGAAATGACATTACATGCG 5654
Qy 1578 -----ThrGlnGlnLeuGlnLeuAlaLysAsnLeu-----AlaIleAl 1591
Db 5655 GTCAAAAGTAATGAGGAAGAAAGAGCTGTACACAAACACTTGTGAGAGAGGTTTTC 5714
Qy 1591 AlaSerAspAsnCysProIleThrGlnGlu----- 1601
Db 5715 TGAAGCTGAATAGACCCGAAATAGAAAGAACTTATGCTGAACATTAGCTCTGCACTACA 5774
Qy 1602 -----LysGluThrSerAlaAspCysValHisPr 1611
Db 5775 AACAGACAGTTGAAAAACCTCTGAAGAACCTAAGTGAAATACAGACTGACCTGAACATGC 5834
Qy 1611 oLeuGluGluLysIleLeuLeuLeuThrGluGluLeuHisGlnLysThrAsnGluGlnGlu 1631
Db 5835 GAAAGTGACACAGACAGAGTGTATGCGTGATTCATTAGACAGAAACAAAGCAACAGA 5894

```

QY 1631 uLysLeu-----LeuHisGluLysAsnGluLeu---G1 1641
 Db 5895 GTCCCTTAAGTGCACAGAGAACTTCGAGAGCCGCTTCATAGAGAGCCAGGCGACAGAA 5954
 QY 1641 uGlnAlaGlnValGluLeuLysCysGluValAlaHisIleuMetLysSerMetIleGlu 1661
 Db 5955 ACACCTAGCTGTGGAGCTCACT---AAGCGTAGGCGGCTGATGATGCTATGCAAGTGA 6011
 QY 1661 rLysSerSerLeuGluSerLeuGlnHisGluLysAsp-----ThrGlu1 1677
 Db 6012 AAAAACCTTTTGAAGGCAAAATTCAGAAAAAATGATATATATAGATGCTGTGGAGA 6071
 QY 1677 nGlnLeuLeuAlaLeuLysGlnMetGlnValValThrGlnGluLysGluLeuGlu 1697
 Db 6072 GGAGCTTTGATGTCACAGTAAACAGGTTGCAAGAAATGAGGACAGACAGACAGATCCA 6131
 QY 1697 nGlnThrHisGluHisLeuThrAlaGluValAspHisLeuLysGluAsnIleGluLeuGlu 1717
 Db 6132 AGAAGAAAGAGAAATTCAGTCCACAGCAAAAGAGCTATGAAA-----GCAGAGCGAGG 6185
 QY 1717 yLeuAsnPhelLysAsnGluAlaGlnGlnLysThrThrLysGluGlnCysIleuLeuAsnGlu 1737
 Db 6186 C-----CCAGTTGAAACAACTTACATACAGGA 6212
 QY 1737 uAsnLysGluLeuGlnSerGlnHisArgLeuGlnCysIleuLeuMetLys 1757
 Db 6213 GACGAGAAATTAATGAGAGGAAAACTGAGAACTGACATGTCACAGTGAAGAAAGTACGGA 6272
 QY 1757 sSerLeuLysAspLysGluSerAlaLeuGluThrLeuLysGluSerGluGlnLysVal11 1777
 Db 6273 TGACCTTCAAAAACAAAGTGAAGCTCTAGAA-----AT 6305
 QY 1777 eAsnLeuAsnGlnGluMetGluMetValMetLeuGlnGluGluLeuLysAsnSerGlu 1797
 Db 6306 AGATGTGAGAGAACAGTCACTAGT---TTTATGAGCTGGAGACAGAAAAATACGGA 6362
 QY 1797 nArgThrValIleAlaGluArgAspGlnLeuGlnAspAspLeuArgLysValGluLeu 1817
 Db 6363 ACTAATGATTTAAGACAGCAAAACCAAGCATTTGAAAAAGCATTTAGAAAAAATGACAAA 6422
 QY 1817 tSerIleGluThrGlnAspAspLeuArgLysAlaGlnGlnAlaLeuGlnLysAsn 1837
 Db 6423 ATTTTAGATGAGCAAGCCATTTGACAGAGAACATGATGAGATGATTTCCACAGC----- 6477
 QY 1837 pLysValGlnGluLeuThrSerGlnIleSerValLeu----- 1849
 Db 6478 -CAATACAGAACTAGAACACCAACTTAAGTTGTTCTCCTGATTCAGCCATCAGTGA 6536
 QY 1850 -----GlnGluLysIleSerLeuLeuGlnGlnMetLeuLysValAlaThrVal 1867
 Db 6537 ACATCAAACTAGAGAGCTTGAACAGCTTGAACAATCATCTGAAAAAAGAAAAAGACAAATG 6596
 QY 1867 lLysGluThrLeuSerGluArgAspAspLeuAsnGlnSerLysGlnHisIlePheSerGlu 1887
 Db 6597 CAGTGTGCTTTTCTCTCTAAAGAGCAGCTTCAAAAGGGATATACAAAGAAAGAAATGAGA 6656
 QY 1887 uIleGluThrLeuSerLeuSerLeuLysGluLysGluPheAlaLeu----- 1902
 Db 6657 AATAGAGAAACAGTGAATCAGAGTAAGAGAACTGAGCGCTCTTGTAGAGGAGCG 6716
 QY 1903 -----GlnGlnAlaGluLysAspLysAlaAspAl 1912
 Db 6717 AAAACACTTGGAGCTGTAGAAAGCTAAACCAAGATTTGCTCAGAGTACATTCAGAGC 6776
 QY 1912 AlaAlaGlyThrIleAspIleThrGluLys---IleSerAsnIleGlnGluLeu 1931
 Db 6777 TGAAGAGATGCCATAGACAGAAAGAAAGAAAGATTAACAATCTAGAGAGCAATTAAGA 6836
 QY 1931 uGlnGlnAlaThrAsnLeuLysGluThrLeuLysGluArgLysLeu---IleGlnCys 1950
 Db 6837 ACAGTTTAGAGAGAACTGGAATAAGATGAAGAGTTCAACAATTTACATATGCAAT 6896

QY 1950 sLysGluGlnLeuAlaLeuAsnThrGluHisLeuArgLysThrLeuLysSerLysAspLe 1970
 Db 6897 AGAAATACAGAAAAAGATTCATACCCGCTACAGAAATCTGAACAGAAACAAAT 6956
 QY 1970 uAlaLeuGluLysMetGlu----- 1976
 Db 6957 ATTTAAGATTCACATGAGAAACTGGGACTGCCATTAAGAAATGATGCCATCTTAC 7016
 QY 1977 -GlnGlnArgAspGlnAlaAlaAsnLysValIleAlaLeuThrGluLysMetSerSerLe 1996
 Db 7017 TCAGAGCCACATGCTCTTGGGAAATTTGCTCAATTAATACAGAAAGAGGTAGA 7076
 QY 1996 uGlnGluGlnIleAsnGluAsnValThrThrLeuLysGluGluGluLysGluThr 2016
 Db 7077 AATTGACCAATTAATGAAACAACTGAGTACGAACTCCAG----- 7113
 QY 2016 rPheLysLeuGlnArgProSerLysGlnIleSerSerGlnMetGluLeuArgL 2036
 Db 7114 -----CACCACTTAAATTTCAACACATTAACAAGTTATTTGAAGAAAAATGCA 7163
 QY 2036 uSerLeuLysThrLysAspLeuGlnLeuGlu----- 2046
 Db 7164 ACTGATTAAGGAGCTTGAACCCAAATAGAAATGTTGATGATGATCAAGAAATGCTGAA 7223
 QY 2047 -----GluAlaGluLysGluIleSerLysAlaThrAsnGluIleLysAsnLeuThrAlaLys 2065
 Db 7224 GAGAAATAGAGAAAGAAATTAAGACGCTCAATGAGAGTGAAGTGAAGAACTTCAACAGCA 7283
 QY 2065 sIleSerSerLeuGlnGluGluIleLeuGlnAsnAlaSerIleLeuAsnGlnAlaValSer 2085
 Db 7284 ATTTGCAATTAATTTGACAGAGAACATCAATGATGCTCATCTCCCTCAGAA----- 7335
 QY 2085 rGluArgLysAsnLeuArgHisSerLysGlnLeuValSerGlu-----LeuGlu 2102
 Db 7336 -GAAGCAGACAGTTTAACATCAATGATGATGTTATGCTGAAGAACTGCCCTTGA 7394
 QY 2102 uGln-----LeuSerLeuThrLeuLysSerArg 2111
 Db 7395 ACACCAAGTGAACCCCTAATGAGAAATGACCTTCATGAATAATGCTTAAAGAAAC 7454
 QY 2111 gAspHisAlaPheAlaGln-----SerLysArgGluLysAspGlu 2124
 Db 7455 CAATTTAAATGAAATGATCAATCACAGAAATTAATTCAGCTTAAAGAGAAAGT---GA 7511
 QY 2124 uAlaValAsnLysIleAlaSerLeuAlaGluGlu-----IleLysIleLe 2139
 Db 7512 AAGTGTGAAGAAATTAAGCATACAGAGATAGCTTAACGTGCTTAATCATCT 7571
 QY 2139 uThrLysGluMetAspGluPhe-----ArgAspSerLysGluSerLeuGlu 2154
 Db 7572 GAGCAAGACAAACCTGAAACATGAGAAATGCTTACAGAGAGATGCTTAAATCCCTAGA 7631
 QY 2154 nGlnGlnSerSerHisLeuSerGluGlu-----LeuLysThrIleLys 2168
 Db 7632 AAATCAGACATCTCAATTTTGAAGAAATGCAAAAGTTCCTCAATTAATTTGGA 7691
 QY 2168 sThrGlnLeuGlnMetLeuLysGlnGlnLysGluAspIleAsnAsnLysLeuAlaGluLys 2188
 Db 7692 AACCAAGTTGCTCAACTTGAAGACACTGTTAGTCAAGAGCACTTAACCTTACCAGCTG 7751
 QY 2188 sValLysGluVal-----AspGluLeuLeuGln----- 2197
 Db 7752 TTATTAACAATTAAGAAAGCATGCAAGAAAGCCAGTTTGAACAGAAATGCTTCAAAA 7811
 QY 2198 -----HisLeuSerSerLeuLysGluGln-----LeuAspGlu 2208
 Db 7812 GAAGATTTGAACCTTACAGAAATTAAGTGAAGAAAGAGTGGCTGCTCTTGTACATGA 7871
 QY 2208 nIleGlnMetGluLeuArgAsnGlu----- 2216
 Db 7872 AATCCAACTTGAAGAGTTCAGGAATATGCAAAATTTCTGCAAGATATACAAATTTTC 7931
 QY 2217 -----LysLeuArgAsnLysGluLeuGlu 2224


```

Oy 2780 -ProvalrThrProGluuArgSerGluMetProSerLeuHisLeuGlySerProIlyrSylsse 2799
Db 9945 TGAGAGTCAGAAAAGAAAGAACTTCACCTAAATCTACTTTGGAAACACAGAAACAACCT 10004
Oy 2799 rGluSerSerThrLyAsrValValSerProAsnArgSerGluIleuLeuSerGlnLeuVa 2819
Db 10005 ACAGAACGAATCCACCAAAAAAATGAAATCAACAGAGAAATGCTATATGATCCACAGTTGTC 10064
Oy 2819 lMetSerProGlyLyThrGlyLeuHisLyHisIleLeuSerProSerLyValGlyLe 2839
Db 10065 AGAAGAAACAAGGTCGAAACTAGACCTCAGGTACTTCTTGAATCTGAGAAAGCTGCAAT 10124
Oy 2839 uHisLyLyAsrGAlaLeuSerProAsnArgSerGluMetProThrGlnHisValIleSe 2859
Db 10125 TCGGGAATAGATAGATACCTGATAGAGAGCGGGA----- 10161
Oy 2859 rProGlyLyThrGlyLeuHisLyAsnLeuThrGluSer-ThrLeuPheAspAsnLeuS 2879
Db 10162 -----TTCCAGCGCACAGCTGCAGAGAGATGATGTACTG6ACAGTCGTC 10205
Oy 2879 erSerProCysLySGlnGln 2885
Db 10206 GCCACCTTCCTCCCTCAGAGG 10225

RESULT 10
US-10-171-311-5
; Sequence 5, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangt
; APPLICANT: Glati, Karen
; APPLICANT: Gannavazapu, Manjula
; APPLICANT: Hoerst, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 12337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12055, 12126, 12288
; OTHER INFORMATION: n = A,T,C or G
US-10-171-311-5

Alignment Scores:
Pred. No.: 3,15e-56 Length: 12337
Score: 1022.50 Matches: 707
Percent Similarity: 36.54% Conservative: 655
Best Local Similarity: 18.97% Mismatches: 1128
Query Match: 6.92% Indels: 1237
Db: 9 Gaps: 153

US-09-150-867-1 (1-2954) x US-10-171-311-5 (1-12337)
Oy 70 ArgSerAlaLeuGlnGlyTyrAsnGlyThrIlePheAlaTyrGlyGlnThrSerSercly 89

```

Db 51 CGATCCCGCACTGAG-----CCGGAGACTGCTTCACCTTGCGCGGGGAGGCCCGCA 104
 Qy 90 LysThrTyrThrMetMetGlyThrProAsnSerLeuGly-----IleIleProGln 106
 Db 105 -----CCGAATCGGCTCTTAGCGCGGTGGAGCTTGCCGCTCC 140
 Qy 107 AlaIleGlnGluValPheLeuYsIleIleGlnGlnIleProAsnArgGlnPheLeuLeuArg 126
 Db 141 CACCTCCCTT-----CCAAATCGACCTTCCCTTCTCTATC 173
 Qy 127 ValSerTyr-----MetGlnIleTyr-AsnGlnThrValYsAspLeuLeuYsAspAs 144
 Db 174 CCGAACCAACCCCTCAACCCCTGTTTCCCTGCTTCCTTGACAGAGCCATGAGAGCA 233
 Qy 144 PArgTArgLysLysProLeuGlnIleArgGlu-----AspPheAsnArgAsnVa 160
 Db 234 GCGAGAGACAGAAAGAAAGCTGGAGCGCGGCAAGCCAAAGCTTGCCCACTTTCGACAAAGAA 293
 Qy 160 I TyrAlaValAspLeuThrArgGlnGluLeuValMetValProGlnHisValIleGlnTrpI 180
 Db 294 AGCTCAGTGGAGTGGCGCAGACTCT-----TC 320
 Qy 180 eLysLysGlyLysLysAsnArgHisTyrGlyGlnThrLysMetAsnAspHisSerLeu 200
 Db 321 CAAGAGACAGAAAAAAGAGAA-----AAACCTCAACGACAGCAAAACAGATGCTC 371
 Qy 200 GserHisThrIlePheArgMetIleValGlnSerArgAspArgAsnAspProThrAsnSe 220
 Db 372 ACCACACACATGATTGATATTGATGTCATACAGCTGAATGAATGATACATAAATGATTC 431
 Qy 220 rGluAsnYsAspGlyAlaValMetValSerHisLeuAsnLeuValAspLeuAlaGlySe 240
 Db 432 TCAGAGAGAGAAATCACTGATGATCCGATCTACAAATATAGAACTTACATAGTGG 491
 Qy 240 rGluAlaArgAlaSerGlnThrArgAlaGlnGlyAlaArgLeuYs----- 254
 Db 492 AGAAATTAACCACTCATGACAGCGGCTCTCTGTGGAACTGGAAAGTGAATTCACACAC 551
 Qy 255 ----GluGlyYsAsn-----IleAsnArgSerLeuPheIleLeuGlyGlnValIleY 271
 Db 552 AGCAGATGACTCTCAGTCTCAGAGGTAAAGTGGTGCAGATTTGTATG----- 597
 Qy 271 sLysLeuSerAspGlyGlnAlaGlyGlyPheIleAsnTyrArgAspSerLysLeuThr 291
 Db 598 -----AGACAGAGAAAGCCTCAAA 617
 Qy 291 gIleLeuGlnAsnSerLeuGlyGlyAsnAlaYsThrValIleLeuCysThrIleThrPr 311
 Db 618 TTATTATTAAGGAAGAAGATTGGT----- 642
 Qy 311 oValSerPheAspGlnThrLeuSerThrLeuGlnPheAlaSerThrAlaYsHisValAr 331
 Db 643 -----GTTTGATGATTTATTC-----GAACAAAGACACCA 674
 Qy 331 gAsnThrProHisValAsnGlnValLeuAspAspGlnAlaLeuLeuYsArgTyrArgY 351
 Db 675 AGCACTCCGACTCATCTAGCATATGGAAGTGAAGTGGCTGGGAAGACAT----- 729
 Qy 351 sGluIleLeuAspLeuYsLysGlnLeuGluAsnLeuGlnSerSerLeuThrLysAl 371
 Db 730 -GAGATTGAAGAGCTAAACAGAGACTGGAAAGAAATGAGGTTCACCTATGGACTGAAG 788
 Qy 371 a-----GlnAlaMetAlaLysGlnIleHis-----ThrGln 381
 Db 789 ACTGCACAGATTACAGAAATTTGAAGCTGCCATTAAACAAAGACATGCAATTAAACCA 848
 Qy 381 nLeuLeuAlaGlnIleLysGlnLeuHisLysLysLysLysLysArgGlyLeuThrHisLeuTh 401
 Db 849 GCTCAGCTGCTAAATTTCACACAGCAAGAAAGAAAGATGAGCAATGACAGAAATTTT 908
 Qy 401 rAsnIleValValAlaSerSerGlnGlnSerGlnGlnAspGlnArgValLysArgLysAr 421

Db 909 AGAGTGG-----ACAGAACAGAGTCACAAATTAACAGATTCAT 947
 QY 421 gATgValThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAlaSerGlyValSerAs 441
 Db 948 TCAGCAATTACAGCGCTAGTGAACCTGTGAGAACAGACACTCATGAGCAGCTGACGA 1007
 QY 441 P-----PheAspMetLeuSerAr 447
 Db 1008 CTTACTACAGCAAGCAACACAGATCTCTCACTCATCAACAGCAGCTTGAGAACACCA 1067
 QY 447 gLeuProGlyAsnPheSerLys----- 454
 Db 1068 CTTATTAGAGATTATACAAAGAAAGAGAACAGACTTCACATGCAATTAAGTTCTTGCA 1127
 QY 455 ----LysAlaLysPheSerAspMetProSerPheProGluIleAspAspSerValLysThr 473
 Db 1128 ACAGAAATTAATTAAGTATATGAATGACACAGTAAATTAAGTAAATTAAGTAAATTAAGA 1187
 QY 473 rGluPheSerAspPheAspAlaLeuSerMetLeuSerSerAsnGlyIleAspAlaG 493
 Db 1188 AGAAATACAGCAAGAAAGAGACATCATGAGAAATTAACACAAATTAATATAGACAGA 1247
 QY 493 U-----TrpAsnLeuAlaSerLysValThrHisArgGluLysThr-----SerIe 508
 Db 1248 AAGAAACTCTTGACTTAAGGATTAATTAACACTGCTGATTAATTAAGTACAGCAAT 1307
 QY 508 uHisGlnSerMetIleAspPheGlyGln-----IleSerAs 520
 Db 1308 ACAGAACAGATTTGTCACAAAGAACACCAAGCAATTAATTAAGTACAGTCTTA 1367
 QY 520 pSerValGlnPheHisAspSerSerLysGlnAsnGlnLeuGlnTrpLysProLysAspSe 540
 Db 1368 TTCTTAAGCAAAAGAAAGACAGCTCTGACAGAAATTAACAGTAAATG----- 1416
 QY 540 rGlyAspMetAlaGluCysArgLys-----AlaSerPheGluLysGluLys 555
 Db 1417 -GGGACAGTCGAGAACTTCAGAGAAATATCAATAAGACAGCCACTGTGAAACTATAT 1475
 QY 555 eThrSer---LeuGlnGlnLeuGlnSerLys---GluGluGluLysArgLysGluLeuVal 573
 Db 1476 AGTACACCAATGAGCAACAAAGAAACACAAAGAAAGTGAACAACTCCGGCAGAGCTGCA 1535
 QY 573 IeGlnSerPheGluLeuLysIleAlaGluLeuGlnGlnLeu-----SerValLysAl 591
 Db 1536 TGAGATGATGGGACAGACATAGTGCATGAATGAACAAATTAATTAAGCAACACATGCG 1595
 QY 591 aLysAsnLeuGluMetValThrAsnSerArg-----GluHisSerIleAsnAlaG 608
 Db 1596 ACAGATGAGAGAAATGAACACGGGCATAGGAGAAATGAGAAATGCTTTAAGGTCATA 1655
 QY 608 uValGlnThrAspValGluLysGluValAlaArgLysGluMetSerValLeuGluLysAspSe 628
 Db 1656 TTCAAATATATACAGTTAATGAAGATCAATTAAG-----TTAATGATGT 1700
 QY 628 rGlyTrpAsnAlaSerAsnSerAspLeuGlnAspSerSerValAspGlyLysArgLysSe 648
 Db 1701 GGCAATTAATGAACGAAATGAATTAATTAATTCACAGTACTACTCTCAAAAGAAACTCAA 1760
 QY 648 rSerSerHisAspGluCysIleGlnHisArgLysMetLeuGlnGlnLysIleValAspLe 668
 Db 1761 GGAAGAACTAGACATTAATTTTGAAGAAAGTGTGCTTACAGAGACACTTGAAACCT 1820
 QY 668 uGluGluPheIleGluAsnLeuAsnLysLysSerGluAsnAspLysGluLysSerSerG 688
 Db 1821 TGTGTAAGAAATGAGCTTTTCAAGGAGACAGATTCAGAGACCTAGACACAACTACTGA 1880
 QY 688 uGlnAspPheMet-----GluSerIleGlnLeuGluCysGluAlaIleMetAl 703
 Db 1881 ACAAGAAAGTAACTTAATGAAGACATTAAGTCCCTTAGACAGTGAAGATTTGAAGC 1940
 QY 703 aGlu----- 704
 Db 1941 TGAATTTGTTCTGCATCTGAATCCAGAAAGAACTAGAATTAATAACATGAACGAGAGT 2000

QY 705 -----LysAlaAsnAlaLeu----- 709
 Db 2001 TACAATTAACAAGATAAACTTGAAATGTTAGAAAAAGAAAGATGTGTGTAGACAG 2060
 QY 709 ----- 709
 Db 2061 AATGGCTGAATCACAGAAAGCTGAATTAAGAGAGCTGAGAACACAGCTTATTATGTC 2120
 QY 710 ----GluGluLeuAlaLeuMetArgAspAsnPheAsp-----AsnIleI 723
 Db 2121 CGAAGAGAGACTTTCACAACTGAGAGAGATTTAGAAATTAAGACATCGAATTAATTTGA 2180
 QY 723 eLeuGluAsnGluThrLeu-----LysArgGluIleAlaAspLeuLysSe 729
 Db 2181 AAAACTTAAGATTAATTTAGCGATTCATTAACAGCAGATGATGTTTACAGATGA 2240
 QY 739 rLeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGluLysGluThrGlnLysG 759
 Db 2241 AATGAGCAAAAGATAGAAACCATGCGATT-----GAAAGGACAAATTTGATTAAC 2291
 QY 759 uHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeuValGluAsnAlaG 779
 Db 2292 TAAAGCAATCAATTAATTTGCAAAATTTCAAGATTAAGATTTACAGCAGCT----- 2346
 QY 779 uMetTrpAsnGluAsnLeuGluGlu-----AspLeuGluThrLysThr 793
 Db 2347 -CTTGTAATTTCAAAAGTCAGACAAATGACTCTTCAATCAATGAACTTCAAAAGAAAT 2405
 QY 793 rLysLeuLeuLysGlu-----GlnGluI 801
 Db 2406 TGAATATCTCAGACAAAGAAAGAAAGGAGTACACTTGAAACAAAGATTCACAGAAAT 2465
 QY 801 eGlnLeu-----AlaGluLeuAlaGlyLysArgAlaAspAsnLeuGlnL 815
 Db 2466 ACACACTTAACAGAAATTTGTAGAAAAACAGATGAAGAAAGAAAGATGATTTCAAGCA 2525
 QY 815 sLysValArgAsnPheAspLeuSerValSer---MetGlyAspSerGluLysLeuGlySe 834
 Db 2526 AAAATTTGCCAAGCTTGAAAGCAGAGATAGCATTTCTTAAGATGAAGAAAGAAACCTTGA 2585
 QY 834 uGluIlePheGlnLeuLysGlnSerLeuSerAspAlaGluAlaVal----- 849
 Db 2586 AGACATGTTGAATATCATCTCTGTAGCCAAAGAAAGAAAGATTAATTTCTTAGACTC 2645
 QY 850 -----ThrArgAspAla-----GlnLysGluCysSerPheLeuArgSerG 863
 Db 2646 CATTAAGTCCAAATCCAAAGACTCTGTGGGAAAGAAATGAATATCTTATAGAGA 2705
 QY 863 uAsn-----LeuGluLeuLysGluLysMetGluAspThrSerAs 876
 Db 2706 AATGAGAGACCTCAAAACAAATGTTATGACGCTTAATGAAGATGAAAGCAAGAAAGAA 2765
 QY 876 nTrpTrpAsnGlnLysGluLysAlaAlaSerLeuPheGluLysGlnLeuGluThrGluL 896
 Db 2766 CACTTTTTCATTTGCGAAAAAACTTTGAAGTAACTATCAAGAGCTTACAAAGAGAGTA 2825
 QY 896 sSerAsnTrpLysLysMetGluAlaAspLeuGln-----LysGluLeuG 911
 Db 2826 TGTGCTCTTTCAGAAAGTAAAGATGATTTAGAGACAGTAAATAATTAACAGGAATTA 2885
 QY 911 n-----SerAlaPheAsnGluIleAsnTrpLysLeuAsn----- 921
 Db 2886 GTATTAAGAACTTAAGCACTTATATGAAGAGCTTCATTTCCAAAGAAATTAATCCAAC 2945
 QY 921 ----- 921
 Db 2946 TACAGTGAATGAAGTTCTGTGTGATGAAGACAAAACTTTGTAGACAGAAACATT 3005
 QY 922 ----GlyLeuLeuAlaGlyLysValProArgAspLeu----- 933
 Db 3006 GGAATATGGTGAGGTTGTGAAGAAATACACAGAACTCATGAAAAAACTTGAGGTTAAC 3065

934 -----SerArgValGluLeuGluLysValLysGluPheSerLysGluLeuGluLys 951
3066 CAACGAGAGAAATTAAGCTGTCACAGACAGCTGTCGATCTTGAACAAATGGAACA 3125
951 salLeuGluGluLysAsnAlaLeuGluAsnGluValThrCysLeuSerLysLys 970
3126 GAAACATGTCAGATTTGTTTCTTAATGAAAGAGTTAAATCTTTAAAGCAAGAGAA 3185
971 -----PheLeuProAsnGluValGluCysLeuLysAsn 981
3186 ACAAGTTTCATTTGAGATGTAGAGCTAGAAATCATTTATACCAACAAGGAGAGAAA 3245
982 -----GlnIleSerLysAlaSerGluGluIleMetLeuLeuLysGlu 995
3246 TGACAGTCATGTATACCTCAAGTAAAGCTCTTTATTTAGATGAGAGTGTGACCATGACAG 3305
995 nGluGly-----GluHisSerAlaSerIleIleSerLys 1006
3306 CAGGGTCTGTAAGATCATGTTCTTAAGTAAATTAAGTTTGTGAGAAATCAAAAT 3365
1007 -----GlnGluIleIleMetGlnGluGlnSerGluGln 1017
3366 AATGTGGAAGATAAATTTCTTTGAAATATGCTGTGAGAGAAATGAAAGTAAAGCA 3425
1018 -----IleLeuGlnLeuThrAspGluValThrHisThrGlnSer-----LysValGlu 1033
3426 ACAGTTGATTTGATGACTTACCATCATGTACACAAAGAAATCATCTTACAGCACTCA 3485
1033 nGlnThrGluGluGlnThrLysLeuGlu-----MetLysLysMetHisAspAs 1048
3486 ACCAAGTGAATGATTAACCTTCAGAAAGAACTCAATGATCTTAATCAGAAACAAATGA 3545
1048 PleuPheGluLysTyrIleArgAsnLysSerGluValGlu 1061
3546 TTTA-----AGGCTACAGATGAGAGCCCAACAGCATTTGCCCTCTCTGTGT 3590
1062 -----AspLeuLeuArgGlu-----MetGlnAsnLeuLysGlyThr 1073
3591 TTATTCACATCATGTGTCAGAGTGTGCTGATATATGAAATGAAAGATTAAGCTCT 3650
1074 -----MetGlnSerValGluValLysIleAlaAspThr 1084
3651 TTGCACTTTAAAGAGAGCTTATTTTCTCAGAGAGAAAGATCAAGAACTTCAGAA 3710
1084 rLysHisGluLeuGlu-----GluThrIleArg-----AspLysGluGlnLe 1098
3711 AATACACGTTTGAATCTACAGCTATGAAACACAAAGAAAGAGTGAAGAAAGCC 3770
1098 uLeuHis-----GluLysLysTyrPhePhe 1106
3771 TTTACATCTGCTCATTTGAAAACTTCAAAAGCAGTGTGGAAGAAATGTTCTTATTTT 3830
1106 eGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProPheSer 1126
3831 ACAGACTTTTTCAGAGTCTC-----CTTGGTGAATTTATATCTCTCTCTT 3875
1126 sLeuValGluGluLysAsnSerGlnAspProIleGluIleAsnAspTyr 1141
3876 AAAATGTGAAGTAAATGCAACAAAGAAAGAAATTCGTGATTTACATTTCTGAAATGA 3935
1142 -----HisAs 1143
3936 AGATCCAGAAATTAACAGATTAATGATTAAGTCAAGACTTTCAGAAAAATATGACAC 3995
1143 nLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr 1160
3996 TCTTCTCAACAAAGTAAACAGAAATTAACAAACCTTGTGTTCTTCAACAGACATAG 4055
1161 -----GluArgAsnSerLe 1165
4056 CAAGATCTGGGACAGACAGATGTATGAAACTTGAATTTGAGAGAAAGAAACCTTCC 4115
1165 uLysGluGlnVal-----IleAs 1171

4116 AAAAGAGAAACAGACTTTTATCATTCATCTCAGATGACCAATTTGGAAGACATTA 4175
1171 PleuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLysSerAspLeuGln 1191
4176 TGTCAATCATTAACAAAGATTAATCTTCTGTCAGATCTGAAATTAACCTTGAAGA 4235
1191 sProLysGlnAspLeuGlu-----GluGlyGlu 1200
4236 ACAAGTTCAAGATTAAGAAAGCTTCATTCCTCTTGCAGCAACAGTTGAAGAAACTGA 4295
1200 uValLysLeuLeuLeuGluMetGluLeuLysGlyHisLeu 1214
4296 ACAAAATATGAGCAGAGATTCACCTGTTACAGAAAGAGCTTCAAGCTGTAGATC 4355
1215 -----ThrAspSerGlnLe 1219
4356 CACGTTCCGCAAGCTTACCTGTTGATTCGGTGTATTAACAGAAATGATGACAGAG 4415
1219 uSerIle-----GlnLysLeuGlnLeuGlnAsnLeuGluValThrGluLysLeuGlu 1221
4416 AACAAATGTACCTCGAAAGTTGTGAAAAAGATTAATGCTACAAATAGAGTTTCTGG 4475
1222 -----GlnLysLeuGlnLeuGlnAsnLeuGluValThrGluLysLeuGlu 1236
4476 TGAATTTGAGTGAAGAGAGAAACAAATATCTTAAGTTGCTTGAAGAAATACATACAG 4535
1236 nThrLeuGlnGluGluMetLysAsnIleThrIle-----GluThr 1249
4536 ACAATTAAGAAAGAAAGTGTATGTTGTCATAGATTAATGATTTGCTTACAA 4595
1249 gaAsnGluLeu----- 1252
4596 AACTGAACCTCTACAGATATCTGGGAAAAAAATCTGCATCATCAACCAAGCACA 4655
1253 -----GlnThrAsnPheGluAspLeuLysAlaGluHisAspSerLe 1266
4656 TGCTGTGTGTCACAGAAACAAATATTTTATGAATGAATTAATTCACAGATCAAAAT 4715
1266 uLysGlnAspLeuSerGluAsnIleGluGlnSerIleGluThrGlnAspGluLeuArgAl 1286
4716 TGTGTTTCAGACTTTTTCAGACAGCTGTGAT-----GTGAATTTAAAGAAATTTTAAAC 4769
1286 alaGlnGluGlnLeuArgGluGln-----LysGlnLeuValAspSerPheArgGln 1303
4770 ACTTAGTAAGAGTTTGAAGAACTGGAAGAAATTTTATTCAAATAGATATCCCA 4829
1304 -----GlnLeuLeuAspCysSerValGlyIleSer-----SerProAs 1316
4830 TGATATACCAAGATCAAAAGCAGCTGTGCTGATATTTCAAGAAATGTTCTTCAAAAG 4889
1316 nHisAspAlaValAlaAsnGln-----GluLysValSerLeuGluValAlaAsn 1333
4890 TAAACATTTATTTAGTTAGACAGCTCTATTCATGATGAAATTTCTACGTGCAAGTATGC 4949
1333 r-----LeuGlnSerGluMetLeuArgGlyLysArgAspGluLeuGlnThr 1348
4950 TTCTAGACAACTAATGTTGATGAAGAAAGCTGTGAGATATGAGACAGAACTTGTAGC 5009
1348 rSerCysLysAlaLeuValSerGluLeuGluLeuArg-----AlaHisValLysSerVa 1367
5010 ACAATACCAAGAACTACACAGGACAGCAAGCAATTTTAAAGCAACATATGCGGCAAT 5069
1367 lGluGluLysLeuGluIleThrLysLysLeuAsnGlyLeuGluLysGluIleLeuGlu 1387
5070 GAG----- 5073
1387 rLysSerGluGlnSerGluValLeuLysSerMetLeuGlnLysLeuLysGluAspAsn 1407
5074 -AGACAGGAGAAAGACAG-----GAAACGCTACAGAAAGAGATTA 5114
1407 nLysLeuLysGluGlnAlaGluLysTyrSerSerLysGluAsnGlnPheSerLeuGluGlu 1427

Db 5115 GAGACTTAATAGCAATTAGCCAGAGATCCTCATAGATATATGAAACCTGGTTTCAGA 5174
 Qy 1427 uValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuLysAlaGluLeu 1447
 Db 5175 G-----AGAGAGAGGGCTTTTAGAGAGCTGGAGACCTAAAGACCTGCTCTTT 5225
 Qy 1447 sAlaIleGluLysArgLeu-----GluIleLysAspArgAspArgPheGluLeuValG 1465
 Db 5226 AGCTGGAGAGAGAGAGAGCTGTGTGTGAGCTGGCAACAGCACT-----ACGCA 5273
 Qy 1465 nThrAlaAsnThrAsnLeuValGluGlyLysLeuGluThrProLeuGlnAlaAspHisG 1485
 Db 5274 AACACAGATGGAATGAAACCAAGAGAACTGAA-----GAACAAACATTTAA 5324
 Qy 1485 uGluAspSerIleAspArgAspSerGluGluMetGluIleLysValLeuGlyGluLys 1505
 Db 5325 AGAAAGAAATTAAGACACAAAACCTGAAGATGTGCTGCTGAGATTTGTCTAATGAA-- 5382
 Qy 1505 uGluArgAsnGlnIleLysLeuGluLysLeuGluGluLysLeuGluLysLeuSerAsn 1525
 Db 5383 -----AGTATGCACTCCAGAAAGCTAAT-----AATAGACT 5414
 Qy 1525 sLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLysAspAsp----- 1542
 Db 5415 TTTGAGATCCTCTTGAAGTTGTAAAGCAACACAGCAGCTGTTGAAGAAACATTTGCTG 5474
 Qy 1543 -----LeuGlnGlnLysLeuGluLysLeuSerLeuSerGluAsnIleIle 1557
 Db 5475 CCATGCTCTTGAGATTCATAGATAGATCTAGTAAAGCCAGTCATCTGCGACGCTAATTTG 5534
 Qy 1557 uLysGluAsnIleAspThrThrLeuLys-----HisHisSerAspThrGluAlaG 1574
 Db 5535 GAGGTCAAGAACAGAGGATCTGTAAAGTCAATGTCATGATGAGAACATACAGAGTTAC 5594
 Qy 1574 nLeuGlnLys----- 1577
 Db 5595 AGATGAATCCATTCCTCTTATTTCTGAAGTGAATGCCAAGAAATGACATTAACATGTG 5654
 Qy 1578 -----ThrGlnGlnGluLeuGlnLeuAlaLysAsnLeu-----AlaIleAl 1591
 Db 5655 GTCAAAAGTAACTGAGAGAGAGAACAGAGCTGTACACAGACTTGTGAGAGAGGTTTGC 5714
 Qy 1591 aAlaSerAspAsnCysProIleThrGlnLys----- 1601
 Db 5715 TGAACATGAATAGACCTGAAAGTGAAGAACTTATGCTGAACATTAGCTCGACTACA 5774
 Qy 1602 -----LysGluThrSerAlaAspCysValHisPr 1611
 Db 5775 AGCAGCACTGTAATACTCTTAGAAGCCATAGTAACTGAACTGAGCTTGAACATGTC 5834
 Qy 1611 oLeuGluGlnLysIleLeuLeuThrGluLeuHisGlnLysThrAsnGluGlnG 1631
 Db 5835 GAAAGTGAACACAGACAGAGTGTATGCTGAGTCAATTAGACAGAAACAGAACAGACA 5894
 Qy 1631 uLysLeu-----LeuHisGluLysAsnGluLeu--G 1641
 Db 5895 GTCCCTTAAGTGCACAAAGAACTTCAGAGAGCGCTCATGAGAGTCCAGGCGCAGAGA 5954
 Qy 1641 uGlnAlaGlnValGluLeuLysCysGluValGluHisLeuMetLysSerMetIleGlu 1661
 Db 5955 ACACTAGCTGTGAGCTCACT--AAGGCTGAGGGCGTCACTGATGCTATGACAGAGA 6011
 Qy 1661 rLysSerSerLeuGluSerLeuGlnHisGluLysHisAsp-----ThrGlnG 1677
 Db 6012 AAAAATCTTTTGAAGGCAAAATTCAGAGAAACATGATATATAGATCGCTTGACACA 6071
 Qy 1677 nGlnLeuLeuAlaLeuLysGlnGlnMetGlnValAlaThrGlnGlnLysGlnLeuG 1697
 Db 6072 GGAGTTGTTATGTGCAAGTACAGAGTTCAGAAATTTGAGAGGAGAGCAACAGACATCCA 6131
 Qy 1697 nGlnThrHisGlnLysLeuThrAlaGluValAlaAspHisLeuLysGlnLysGlnLeuG 1717
 Db 6132 AGAAGAAAGAAATTAAGTCTGTCAGACAAAGAAAGCTATGAAA-----GCAGAGCGAG 6185

Qy 1717 rLeuAsnPheLysAsnGluAlaGlnGlnLysThrThrLysGluGlnCysLeuLeuAsnG 1737
 Db 6186 C-----CCAGTTGAACACAACTTACTACAGCA 6212
 Qy 1737 uAsnLysGluLeuGlnSerGlnHisArgLeuGlnCysGluIleGluLeuMetLys 1757
 Db 6213 GACAGAAAAATTAATGAGAGAAAAAACTAGACATCTCAAGCTGAAAAAGATACGTA 6272
 Qy 1757 sSerLeuLysAspLysGluSerAlaLeuGluThrLeuLysGluSerGluGlnLysVal 1777
 Db 6273 TCACCTTCAAAAACAAAGTGAAGCTCTAGAA-----AT 6305
 Qy 1777 eAsnLeuAsnGlnGlnMetGluMetValMetLeuGlnGluLeuLysAsnSer 1797
 Db 6306 ACATGTGGAAGAACAGTCAATGAG--TTATAGAGCTGGAACAGAAAAAATATCTGA 6362
 Qy 1797 nArgThrValIleAlaGluArgAspGluLeuGlnAspLeuArgGluSerValGluMe 1817
 Db 6363 ACTAATG----- 6369
 Qy 1817 tSerIleGluThrGlnAspAspLeuArgLysAlaGlnGluAlaLeuGlnGlnLysAs 1837
 Db 6370 -----GATTTAAGACAGCAAAACCAAGCATTTGAAACAGTATAGA 6410
 Qy 1837 pLysValGlnGluLeuThrSerGlnIleSer-----ValIle 1849
 Db 6411 AAAAATGAGAAAAATTTTATAGATGAGCAAGCCATGACAGAACATGACAGATGATT 6470
 Qy 1849 uGlnGluLysIleSerLeuLeuGlnLysGlnMetLeuThrAsnValAlaThrValLysG 1869
 Db 6471 CCACAGGAATATACAGAAACTGAAACAGCACTT-----AAGTGTGCTCCGATTCGA 6524
 Qy 1869 uThrLeuSerGluArgAspAspLeuAsnGlnSerLysGlnHisLeuPheSerGluIleG 1889
 Db 6525 GCCTATCACTGAA-----CATCAAACTGAGAGAGTTGA 6557
 Qy 1889 uThrLeuSerLeuSerLeuLysGluLysGlu-----PheAlaLe 1902
 Db 6558 ACAGTTAGCAATTCATCTCAAAAAGAAAAACACAAATTCAGTACGCTTCTCTTAA 6617
 Qy 1902 uGlnGlnAlaGluLysAspLysAlaAspAlaAlaArgLysThrIleAspIleThrGlu 1922
 Db 6618 AGACGACCTCAAGAGATATCAAGAAAGAAATGAGAAATGAGAACTGAGATTGAC 6677
 Qy 1922 sIleSerAsnIleGluGlnGlnLeuGlnGlnAlaThrAsnLeuLysGluThrLeu 1942
 Db 6678 ACTAAGAGAACTGAGACAGCGCTTCTGTGAGTGCAGATCTTCAAAAGTAGAGCA 6737
 Qy 1942 rGluArgGlu--SerLeuIleGlnCysLysGlnGlnLeuAlaLeuAsnThr-----G 1959
 Db 6738 CCGAAACACTTTGGAGCTGTGAGAACGCTTAACACAGATTTGCTCCTAGAAATGCA 6797
 Qy 1959 uHisLeuArgGluThrLeuLysSerLysAspLeuAlaLeuGluLysMetGluGlnGlu-- 1978
 Db 6798 GGCTGAACAGATGCCATGACAGAAAGAAAGAAAGATTAACAATCTAGAACACAAT 6857
 Qy 1979 -----ArgAspGluAlaAlaAsnLysValIleAlaLeuThrGlnLysMetSerSe 1995
 Db 6858 AGAACAGTTTAAAGAACTGGAATAATTAAGAAAGTCAACATTAATCATATGCA 6917
 Qy 1995 rLeuGlnGlnGlnIleAsnGlnAsnValThrThrLeuLysGlnGlnGlnGlnGlnG 2015
 Db 6918 ATTAAATAATCAAGAAAAAGAAATCTACTACCGCTTACAGAACTTGAACAGAAAAACA 6977
 Qy 2015 uThrPheThrLeuGlnArgProSerLysGlnGlnSerSerGlnMetGlnGluLeuAr 2035
 Db 6978 ATTATTT-----AAGATGACATGAGAAACTGGG 7007
 Qy 2035 gLysSerLeuLysThrLysAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 2049
 Db 7008 ACTTCATTAAGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7067

QY 2050 -----LysGluIleSerGluIleThrAsnGluIleLysAsnLeuThrAlaLysIleSe 2067
 Db 7068 ATTTGCTCAATATACAGAGAAAAGAGTAGAAATGACCAATTAATGACAGAGCTTAC 7127
 QY 2067 rSerLeuGluGluGluIleLeuGlnAsnAlaSerIleLeuAsnGluAlaLysSerGluTr 2087
 Db 7128 GAAATCTCCACAGCAACTT-----AAATTTCAACAGATACAGAGGTATTTCAGAA-- 7179
 QY 2087 gGluAsnLeuAlaGHisSerLysGlnGlnLeuValSerGluLeuGlnGlnLeuSerLeuTh 2107
 Db 7180 -----AAAAATGAACTGATTAAGGATCTTGAAACCCAAATAGAAATG 7220
 QY 2107 rLeuLysSerArgAspHisAlaPheAlaGlnSerLysArgGluLysAspGluAlaValAs 2127
 Db 7221 TTTGATGAGTGATCAAGAAATGTGTGAAGAGAAATAGAGAAAGAAATATAGAGAGCTCA 7280
 QY 2127 nLysIleAla----- 2130
 Db 7281 TGAAGTGATTTGAAAACTTCAACAGAGAAATTTGGACAGAGAAAGATCAATGAA 7340
 QY 2131 -----SerLeuAlaGluGluIleLysIleLeuThrLysGluMetAspGluPheArgAs 2148
 Db 7341 TGCTCATTCCTCTCAGAGAGACAGACAGTTTAAACATCAATTTGGATGTATAGC 7400
 QY 2148 pSerLysGluSerLeuGlnGlnGlnGlnSerHisLeuSerGluGluLeuGlnGlnGlnGln 2168
 Db 7401 TGAAGAGCTGGCTTGAAACAGCAAGTAGAAACCGCTATAGAAATAGACCTTCATGAA 7460
 QY 2168 sthr-----GluLeuGlnMetIle 2174
 Db 7461 AAATGACTTAAGAAACCAATTTAAATGATCAGCTAACAGAGAAATTTATTCAGCTT 7520
 QY 2174 uLysGlnGlnLysGluAspIleAsnAsnLysLeuAlaGluIleValLysGluValAspGlu 2194
 Db 7521 AAAGAGAGAGAGCTGAAAGCTGTG-----GAAAGATTTCAACCATCAACCA 7565
 QY 2194 u-----LeuLeuGlnHisLeuSerSerLeuLysGlnGlnLeuAspGlnIle 2209
 Db 7566 GAATAGTGTAACTGAGCTGATATAGATCATCTGACAAAGACCAATCTTGAAGAGT 7625
 QY 2209 eGlnMetGlu-----LeuArgAsnGluLys-----LeuArgAsnIleArgGluLe 2223
 Db 7626 CCTTACAGAGAGATGCTTTAAATCCCTAGAAATCAAGACATCAATCTTGTGAAGA 7685
 QY 2223 uCysGluLysMetAspIleLeuGluLysGluIleSerValLeuArgLeu----- 2239
 Db 7686 AAATGGCAAGGTTCATATATATTTGGAAACAGAGTGTGCTACAACTTGAGAGCACTGT 7745
 QY 2240 -----MetGlnAsnGlu 2243
 Db 7746 TAGTGCAAGAGACTTAGAATTAACCACTGTTATTAACAAATTAAGACATGCAAGAA 7805
 QY 2243 uProGlnGlnGluGluAspAspValAlaGluArgMetAspIleLeuGlnSerArgAsnGlu 2263
 Db 7806 AGGCCAGCTTTGAA-----ACAGAAATGCTTCAAAAGAAATATGT 7844
 QY 2263 nGluIleGlnGluLeuMet-----GluLysIleSerAlaValIleYrSerGlnGlnHisThrLe 2282
 Db 7845 AAACCTACAGAAATATGTTGAAGAAAAGTGTGCTGCT----- 7881
 QY 2282 uLeuSerSerLeuSerSerGluLeuGlnLysGluThr-----GluAlaHisLysHisCy 2300
 Db 7882 -----GCTCTTGTTCAGTCAATCAATCACTTGAAGCAAGTTCAGGAATATGCAAAATTTCTG 7934
 QY 2300 smetLeuAsn----- 2303
 Db 7935 TCAGATTAATCAACAATTTTCATCAGAACTGAAAGACAAATATTTCAGAAATTAATCA 7994
 QY 2304 -lleLysGlu-----SerLeuSerSerThrLeuSerArgSerPheGlySerle 2319
 Db 7995 ACAGAAAGAGAGATGAGTGGGCTCAGATATATCAGCATTAACCTTGAGAAATTCAGAAAT 8054
 QY 2319 uGlnThrGluHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsnLysPheLysVal 2339

Db 8055 ACAGAGCCAGGTGTGTGAATATCATCATAGTGTG-----ATTTTAGAAAGAAAGCAAGT 8108
 QY 2339 lValIleArgThrAlaAlaValLysGluAspHisSerLeuIleLysAspIleArgGluLysAs 2359
 Db 8109 AGAAATTCGACGAAAAAAATGTTTGTAGAAAAAGAAAGAGCTGTAGAACTACAGAGAGCT 8168
 QY 2359 pleuAlaAlaGluGlnLysArgHisAspGluLeuArgLeuGlnGlnGlnGlnGlnGlnGln 2379
 Db 8169 ATTGGAGGCCAATGAGAAAAACAGAGAGAGAAAGAA----- 8205
 QY 2379 nHisGlyArgLysTrpSerAspSerAlaSerGluGluLeuLys----- 2393
 Db 8206 -----AGAAAGAGAGGCCCTCAGATGTGTGAAGTCTCAAGACCACTAGACTAGCTATT 8258
 QY 2394 -----PheCysGluIleGluPheLeu----- 2400
 Db 8259 TCATAGCAATGAAGAAAGTGATTTTATATGAACTCGAGCTTTAGAGCTGAATCACT 8318
 QY 2401 -----AsnGluLe 2403
 Db 8319 GGCTACCAAGACAGACTTCCAGCTTATTAAGAAAAAGCTGAAACCTTCAAGAAAGAGCT 8378
 QY 2403 uLeuPheLysLysAlaAsnIleIleGlnSerValGlnAspAspPheSerGluValGlnVal 2423
 Db 8379 TTTGGTAAAGAAACAAAT--ATGACATCTCTTCAGAAAGACTTAAAGCCATTAAGGA 8435
 QY 2423 lPheLeuAsnGlnValGlySerThrLeuGln-----GluGlu 2435
 Db 8436 TCACCTCGCAGAGAGCAAGAAAGAAATTTGCCATTTTAAGAAAAAGAGATGAGAGTGAAGT 8495
 QY 2435 uLeuGlnHisLysLysGly----- 2441
 Db 8496 ACAGAGAGAGCAAGAAAGCCGTGATGTTGAGCCACTTCTATTAAGAACTGATAGAGCAT 8555
 QY 2441 ----- 2441
 Db 8556 TGCACTCCAGACAGATGGAGCTGAAAGATCAGTACACACATCACTCCACAAATTTCT 8615
 QY 2441 ----- 2441
 Db 8616 TGTTAAATATGACAGATACAAATTAATTTACAGAGTGAATGTTCTCCAGAGAAAGCTTAC 8675
 QY 2442 -----PheMetGlnThrLeuGlnGluGluPheGlyAspLeuHisValAspAla 2456
 Db 8676 TGAATATATAGCTCAGCTTACTGAAAAATTTGAGAAAGATGCCAAGAACTACATCT-- 8730
 QY 2456 aLysLysLeuSerGlnGluMetGlnGlnGlnGlnAsnArgArgIleAlaSerThr----- 2473
 Db 8731 -----GCTGAATTTTGGACATGAAATCCAGACATATTTTCAGAAACTGAAACCTT 8780
 QY 2474 -----IleGlnLeuLeuThrLysArg-----LeuLysAla 2483
 Db 8781 AAAGAGGAGAACATATATGTCCTGCTTCACTGAAAGAAATGTGTGATCTTGAAGGC 8840
 QY 2483 aValValGln-----SerLysIle----- 2489
 Db 8841 AGTATATACAGTGTGAGAGATTAAGAGGATCTCATAATTCGAGACTACACATTTCTGA 8900
 QY 2490 -----GlnArgGluLeuThr-----Val 2495
 Db 8901 TGCTTACAGACTAGAGAAATATGCTCCAGTATTCGTATGATCAAGACTGGGGCTAGAGGAT 8960
 QY 2495 lTyrlLeuAsnGln-----PheGluAlaLysLeuGlnGluLysLysGlnGlnAsn-- 2511
 Db 8961 TTATCTTACACACAGCTAGAGGATTTGACATPAGCATCAAGAAAGCCGAGAGAAAGAAAGT 9020
 QY 2512 -----LysGluLeuMetArgArgMetGluHisI 2521
 Db 9021 AAGTCAGACAGATTCCTTCCAAAGAAATTAAGAGATTTACTGAGAGCTGTGCATATGA 9080
 QY 2521 sGlyProSerAla-----SerVal 2527

```

Db 9081 AGCATGACAGGTCTTCTCTCAGTCTCCCTATAGTAGAGAGACCATTTCTAT 9140
Qy 2527 lmetglu-----glugluasnalaargyleuuglylleuylserh 2541
Db 9141 TCAGCAGGTTTCAGACCTTGCGCTAGAGAGAGAAAGCTTACATCAATCTCATC 9200
Qy 2541 rValGlnaspGluSerLysLysLeuGlnSerArgLleuMetLeuGlnGluLeuAs 2561
Db 9201 TCTAAGGATTTAATTACAAAGATGCAA-----CTGCAAAAGAGAACCGCA 9245
Qy 2561 nleuValLysaspAspAlaMetHis-----LysGlyGlyLysValAl 2575
Db 9246 GGTATTATGATGATTCTCAATCTCAGAGAGCTTCAGACTGCGAGGTGAACTACTGCT 9305
Qy 2575 alleuGlnaspLysLeuLeu----- 2582
Db 9306 TGCCCTTCACACAGTTTCTTATGAGAGACCTATGTTTACTAGACAGATTTCGACGGA 9365
Qy 2583 -----SerArgasnAlaGluAlaGluLeuAsnAlaMetGlnValLysLe 2597
Db 9366 GCTGACAGCTAGTACTAGTACAGATGAGCTGGTTTACTAAACCTTTGGAAACAGAAAT 9425
Qy 2597 uThrLysLysGlnAspAsnLeuGlnAlaAlaMetLys----- 2609
Db 9426 ACAAGAACAGGGTGTGAAATATCAGACAGCTATGGAAATGCCCTCCAGAAAGCAGATAGAAG 9485
Qy 2610 -----GluLleGluAsnLeuGlnLysMetValAlaLysGlyAlaValProty 2625
Db 9486 GAGTTTGTATTCTGAATATTCAGGCACTGCATGCACAAATGAATGAGTAAATTAATCTCT 9545
Qy 2625 rLysGlnGluLleAsp-----AsnLeuLysTh 2634
Db 9546 GAAAGAGAACACAGAGAGTGAAGAAACCAAGCAAGACTCTTGCAATATATATACAGCA 9605
Qy 2634 rLys-----ValValLysLleGluMetGluLysLleLysLysLysLysAl 2649
Db 9606 GAACAGACTCTCAAAATGCTGAGATGTCAGAGTGAAGTCAAGTGAAGAA--GACAGAGC 9662
Qy 2649 aThrAspGlnGluLleAlaLysLysSerCysLeuGlnAspLysGlnGluLysLeuAr 2669
Db 9663 AACGGAA-----CTGCAGAGACAGCTGAGTCTGCAAAATGCTGTGC 9707
Qy 2669 garLeuLysGlnGluLeuArGArGAlaGlnAlaAspAsnAspThrThrValCysValPr 2689
Db 9708 TGAAGTGAAGAGTGAAGTTCGCAAACTAAATGGAAGTGAAGAACACATCCAAAGCACA 9767
Qy 2689 oLysAspArgLysLysAlaSerThrProValThrCysGlyGlyLysSerGlyLleVa 2709
Db 9768 GCATTAACACCTAATAA----- 9783
Qy 2709 lGlnSerThrAlaMetLeuValLeuGlnSerGluLysAlaLalaLeuGluArGluLeuSe 2729
Db 9784 -----GAATTCGAGGCTTTTCAGGTTGGAAGTTTAA 9812
Qy 2729 rHisTyrLysLysLysTyrHisHisLeuSerArgThrMetSerSerSerGluAspArg-- 2748
Db 9813 AGATTAAGACAGATGAAGATCTTGCTTAATGACACATTTAGACAGTGAACGAATAATC 9872
Qy 2749 -----LysLysThrLysAlaLysSerAspAlaHisSerSerHis 2761
Db 9873 AAGAGAGCTCCAGTGGCTTTGGAGAAAGAGAAAGCCCAAG----- 9912
Qy 2761 sThrGlySerSerHisArgGlySerProHisLysThrGlnThrGlyArgHisGly--Pr 2780
Db 9913 -TTGGGAGCGCGTGAAGACCGGATTAAGAAAGAACTTGAAGATCGAAGTTTTCACCTTGA 9971
Qy 2780 ovalThrProGluArgSerGluMetProSerLeuHisLysLeuLysSerProLysLysSerG1 2800
Db 9972 GAGTCAGAAACAAAGAAATCTTCAGCTAAATCTTCTTGGACACAGCAACAAACTACT 10031
Qy 2800 uSerSerThrLysArgValAlaSerProAsnArgSerGluLleTyrSerGlnLeuValMe 2820
Db 10032 GAACGAATCCCGAGAAATAAGATACACAGAGATGCTATATGATGCCAGTTGTCTAGA 10091

```

```

Qy 2820 tSerProGlyLysThrGlnGlyMetHisLysHisLysLeuSerProSerLysValGlyLeuH1 2840
Db 10092 ACAGACAGGTGCGAAACCTTAGAGACTTTCAGTACTTCTTCATGATGGAAGAACTTCGAATTCG 10151
Qy 2840 sLysLysArgAlaLeuSerProAsnArgSerGluMetProThrGlnHisValLleSerPr 2860
Db 10152 GGAAATGAGTACTACCTCATGATAGGAGCGGGA----- 10185
Qy 2860 oGlyLysThrGlyLysHisLysAsnLeuThrGluSer-ThrLeuPheAspAsnLeuSerS 2880
Db 10186 -----TTGCACGACAGCTGCGACAGACAGTGAATGATGTAAGTGCAGACTTCGCGCC 10232
Qy 2880 eProCysLysGlnGln 2885
Db 10233 ACCCTTGCCCTCAGAGG 10249

RESULT 11
US-10-171-311-1
; Sequence 1, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumel
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersb, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12462
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-171-311-1

Alignment Scores:
Pred. No.: 3,19e-56 Length: 12462
Score: 1022.50 Matches: 707
Percent Similarity: 36.54% Conservative: 655
Best Local Similarity: 18.97% Mismatches: 1128
Query Match: 6.92% Indels: 1237
DB: 9 Gaps: 153

US-09-150-867-1 (1-2954) x US-10-171-311-1 (1-12462)
Qy 70 ArgSerAlaLeuGlnGlyTyrAsnGlyThrIlePheAlaTyrGlyGlnThrSerSerGly 89
Db 51 CGATCGGCGACGAGAG-----CGCGGAGACCTGCTTCACATTCGGCGGGGAGGCCCCGGA 104
Qy 90 LysThrTyrThrMetMetGlyThrProAsnSerLeuGly-----IleLeuProGln 106
Db 105 -----CCGAATCGGCTCTCTAGCGCCGTGAGCTTGGCGGTC 140
Qy 107 AlaIleGlnGluValPheLysIleIleGlnGluLleProAsnArgGluPheLeuLysArg 126
Db 141 CACCTCCGT-----CCAAATCGACCTTCTCTTATATC 173
Qy 127 ValSerTyr-----MetGluLleTyr-AsnGluThrValLysAspLeuLeuCysAspAs 144

```

```

Db 174 CCCAACCCCTCACCCCTGTTTCCCTGCCCTTTCGACAGAGCCATGGACGACGA 233
QY 144 PARGARGLYSLYSProleuGluIleArgGlu-----AspPheAsnArgAsnVal 160
Db 234 GGAGAGACAGAAAGAGTGGAGCGCGCAAGACCTTGCCACGATTTGCAAGAAAGA 293
QY 160 ILYTVAlAlAspLeuThrIleGluIleValMetValProGluHisValIleGlnTrpI 180
Db 294 AGCTCACTCGGATGGCAGAGTCT-----TC 320
QY 180 eLYSLYSGLYIuLysAsnArgHisTyrGlyGluThrLysMetAsnAspHisSerSer 200
Db 321 CAAAGACAGAAAAAAGAGA-----AAACGTCAAGCAGTAACATGATGTGTC 371
QY 200 gSerHisThrIlePheArgMetIleValGluSerArgAspArgAsnAspProThrAsn 220
Db 372 AGCACACCATGATTTGATTCATCAATCAACAGTCAATGAATGTACATTAATAGTTC 431
QY 220 rGluAsnGlyAspGlyAlaValMetValSerHisLeuAsnLeuValAspLeuAlaGly 240
Db 432 TCAGAGAGTACATCAACTGTGATTCCTGATCTACATATATGAACTCTACACTAGTG 491
QY 240 rGluArgAlaSerGlnThrIleAlaGluGlyValArgLeuLys-----254
Db 492 AGAATATACCATCATGACAGAGGCTTCTGTGAACTGGAAGTGAATTCACACAC 551
QY 255 ----GluGlyAsn-----IleAsnArgSerLeuPheIleLeuGlyGlnValIle 271
Db 552 AGCAGATGACAGTTCAGAGTAAAGTTCGAGTTTGTGTGATG-----597
QY 271 sLYSLYSerAspGlyGlnAlaGlyGlyPheIleAsnTyrArgAspSerLysLeuThr 291
Db 598 -----AGAACAGAAAGCCTCAAA 617
QY 291 gIleLeuGlnAsnSerLeuGlyGlyAsnAlaLysThrValIleIleGlyThrIleTrp 311
Db 618 TTTATTTAAGGAGAGAAATTTGTT-----642
QY 311 oValSerPheAspGluThrLeuSerThrLeuGlnPheAlaSerThrAlaLysValAr 331
Db 643 -----GTTATGATTTTATTCT-----GAAACAAGACGACCA 674
QY 331 gAsnThrProHisValAsnGluValLeuAspAspGluAlaLeuLeuLysArgTyrArg 351
Db 675 AGACAGTCCGACTCATAGAGATGATGAAAGTGAAGTGGCTGGAGAGCAGCAT-----729
QY 351 sGluIleLeuAspLeuLysGlnLeuGlnAsnLeuGluLeuSerSerGluThrLysAl 371
Db 730 -GAGATTGAGAGCTAAACAGAGAGCTGGAAGAAATGAGCGTTACCTATGGGACTGAAG 788
QY 371 a-----GlnAlaMetAlaLysGluLysHis-----ThrG 381
Db 789 ACTGACAGAGTTACAAGAAATTTGAAGTGCATTAAACAAGACATGCAATTATACCA 848
QY 381 nLeuLeuAlaGluIleLysGlnLeuHisLysGlnArgLysAspArgIleThrHisLeuTh 401
Db 849 GCTCACTCTATTTTACACACAGCAAGAAAGATGAGAAATGAGAAATTTT 908
QY 401 rAsnIleValAlaLysSerGlnLysSerGlnLysAsnArgValLysArgLysAr 421
Db 909 AGAGTTG-----ACAGAACAGAGTCAAAAATTTACAGATTCATAT 947
QY 421 gArgValIleThrTrpAlaProGlyLysIleGlnAsnSerLeuHisAlaSerGlyValSer 441
Db 948 TCAGCAATATACAGCTAGTGAAGCTGAGAAACAGACACTATATACACAGCTGACA 1007
QY 441 P-----PheAspMetLeuSerAr 447
Db 1008 CTTTACTACAGCCAAACACAGATCTCATCTATCAACAGCAGCTTGAGAAACAGACA 1067
QY 447 gIeuProGluAsnPheSerLys-----454

```

```

Db 1068 CTTATTAGAAGATTATACAAAAAGAAAGACCTTCACAAATGATTGTTCTTCA 1127
QY 455 ----LysAlaLysPheSerAspMetProSerPheProGluIleAspSerValLysTh 473
Db 1128 AGAGAAATTTAAATGATATGAAATGACACAGATTAATAAAGTAAAGTCAATTAAGA 1187
QY 473 rGluPheSerAspPheAspAlaLeuSerMetLysSerAsnGlyIleAspAlaG 493
Db 1188 AGAAATACAGAAAGAGACATCATGAAATTAACAAAAAATTAATGAAAGAGA 1247
QY 493 u-----TrpAsnLeuAlaSerLysValIleHisArgGlyLysThr-----SerLe 508
Db 1248 AAAGAAACTCTTGACCTTAAGGATTAATTAACACTGCTGATTAATTTCTAGAGAAAT 1307
QY 508 uHisGlnSerMetIleAspPheGlyLys-----IleSerAs 520
Db 1308 ACAAGAACAGATGTGTCCAAAGAACACAGAAATTAATAACATGAAATTAGAGCTGACTA 1367
QY 520 pSerValGlnPheHisAspSerSerLysGluAsnGlnLeuGlnIleTyrLeuProLysAsp 540
Db 1368 TTTCTAAGCAAAAAGAAAGACAGCTTCTGAAAGAAATTAACAGTTAATG-----1416
QY 540 rGlyAspMetAlaGluCysArgLys-----AlaSerPheGluLysGluI 555
Db 1417 -GGACAGTCAAGAACTTCAGAAAGAAATCATTAAGACAGCCAGTTGGAACATGATAT 1475
QY 555 eThrSer----LeuGlnGlnLeuGlnSerLys----GluGluLysLysGluLeuVal 573
Db 1476 AGTACAACGAAATGCAACAAACAAAGAAAGTGAACAACTCCGGGCGAGCTGCA 1535
QY 573 lGlnSerPheGluLeuLysIleAlaGluLeuGlnGluGlnLeu-----SerValLysAl 591
Db 1536 TGAAGATGATGGCAGCAGATCTGCAAAATGAAACAAAGATTAATTAAGACACACATGGC 1595
QY 591 aLYSLYSLeuGluMetValThrAsnSerArg-----GluHisSerIleAsnAlaG 608
Db 1596 ACAGATGAGAGAAATGAAACACAGGCAATAGGAGAAATGAGAAATGCTTAAGCTCANA 1655
QY 608 uValGlnThrAspValGluLysGluValAlaArgLysGluMetSerValLeuLysAsp 628
Db 1656 TTCAAATATTACAGTTATATAGATCAGATCAAGTAAAG-----TTAATGAATGT 1700
QY 628 rGlyTyrAsnAlaSerAsnSerAspLeuGlnAspSerSerValAspLysArgLeuSe 648
Db 1701 GGCATTAATTAAGTAAATTTGCAAGATACATCTCCAAAGAAAGAAACTCA 1760
QY 648 rSerSerHisAspGluCysIleGluHisArgLysMetLeuGluGlnLysIleValAsp 668
Db 1761 GGAAGACTAGGACTAATTTTACAGAAAGAGTGTCTCAGAGACAGCTTGAAGCCT 1820
QY 668 uGluGluPheIleGluAsnLeuAsnLysLysSerGluAsnAspLysGlnLysSerSerG 688
Db 1821 TGTGGAAGAAATGAGCTTTTCAAGGAGACAGATTCACAGAGCTACAGACATATAGCTGA 1880
QY 688 uGlnAspPheMet-----GluSerIleGlnLeuCysGluAlaIleMetAl 703
Db 1881 ACAAGAAAGTAAACTTAATGAAACACATATAGTCCCTAGTACAGTGGAAATTTGAAGC 1940
QY 703 aglu-----704
Db 1941 TGAAGATTGTTTTCATCTGATTCAGAACAGAAAGAACTTGAAATTAACATGACAGAGT 2000
QY 705 -----LysAlaAsnAlaLeu-----709
Db 2001 TACAATTTACAAGATTAACACTTGAAATGTTAGAAAAAGAAAGAAATCTGTGTAGACAG 2060
QY 709 -----709
Db 2061 AATGCTGAATCACAGAGAGCTGAATTAGAGAGGCTGAGAACACAGCTTCTATTATGCTCA 2120
QY 710 ----GluGluLeuAlaLeuMetArgAspAsnPheAsp-----AsnIleI 723
Db 2121 CGAAGAAAGAGCTTCCAAACTGAAAGAGAGATTGAAATTTGACATGCAATTAATTTGA 2180

```

Oy 723 elcuglunsgluthrleu-----LysArgLulLeuIleAspLeuGluArgse 739
 Db 2181 AAACTTAACATATATTAGGCACTTCACTATTAACAGAGATAGCTGTTTACAGAAAGA 2240
 Oy 739 rleuylsglunsgluthrAsnGluPheGluIleuGluLysGluThrGlnLysG1 759
 Db 2241 AATGAGTCAAAAAGATAGAACCATGCACTT-----GAAAGGACAAATTTGATAAC 2291
 Oy 759 uHISGluIleGlnLeuIleHISGluIleGlySerLeuLysLysLeuValGluAsnAlaG1 779
 Db 2292 TAACGAGATCAATTAATTTTGAAATTTCAAGCTAAAGATTACGCGAGTCT----- 2346
 Oy 779 uMetLysAsnGlnAsnLeuGlu-----AspLeuGluThrLysThr 793
 Db 2347 -CTTGTAATTTCAAAAGTCAGAAAGAAATGACTTCTTCAAAATCAATGAACTTCAAAAGAAAT 2405
 Oy 793 rLysLeuLeuLysGlu-----GlnGluI1 801
 Db 2406 TGAATACTCAGACAGAAAGAAAGAGGACTTGTACAGAACAGAAATTCAGAAAT 2465
 Oy 801 eGlnLeu-----AlaGluLeuArgLysArgAlaAspAsnLeuGlnLys 815
 Db 2466 ACAACTTAAACAGAAATTTGTAGAAAAACAGATGAGAGAAAAAGAAATGATCTTCAAGA 2525
 Oy 815 sLysValArgAsnPheAspLeuSerValSer---MetGlyAspSerGluLysLeuGlyG1 834
 Db 2526 AAAATTTGCACAACTTGACGACAGAAATAGCACTTCTTAAAGCTGAAAGAAACCCCTTGA 2585
 Oy 834 uGluIlePheGlnLeuLysGlnSerLeuSerAspAlaGluAlaVal----- 849
 Db 2586 AGACATGTGTAATTAATCTCCTGTTAGCCAGAAAGAAATGATGATTTCTTAGACTC 2645
 Oy 850 -----ThrArgAspAla-----GlnLysGluLysSerPheLeuArgSerG1 863
 Db 2646 CATTAAGTCCAAATCCAAAGACTCTGTGTGGAAAAAGAAATAGAAATACCTTAAGAGA 2705
 Oy 863 uAsn-----LeuGluLeuLysGluLysMetGluAspThrSerAs 876
 Db 2706 AAATGAGAGACCTCAACACAAATGATTTACACTTAATGAAGATGAGAAAGCAAGAAAGA 2765
 Oy 876 nTrpTrAsnGlnLysGluLysAlaAlaSerLeuPheGlnLysGlnLeuGluThrGlnLys 896
 Db 2766 CACTTTTCATTTGCTGCAAAAAAATTTGAAGTTAACTATCAAGAGTTACAAAGAGAGTA 2825
 Oy 896 sSerAsnTrpLysLysMetGluAlaAspLeuGln-----LysGluLeuG1 911
 Db 2826 TGCCTTGCCTCTCAAAAGTAAAGATGATTTAGAAAGACAGTAAATAATAACAGAAATAGA 2885
 Oy 911 n-----SerAlaPheAsnGluIleAsnTrpLysAsn----- 921
 Db 2886 GTATTAAGTAACTTAAGACACTTAATGAAGAGCTTCATTTGCAAGAAATTAATCCAAC 2945
 Oy 921 ----- 921
 Db 2946 TACAGTGAATAATGAAGAGTCTGCTTTGATGAGAGACAAACTTTTGTAGCAGAAACAT 3005
 Oy 922 -----GlyLeuLeuAlaGlyLysValProArgAspLeu----- 933
 Db 3006 GGAATGGGTGAGTGTGTTGAAGAGATACAAAGAACTCATGTGAAAACTTGAGGTAAAC 3065
 Oy 934 -----SerArgValGluLeuGluLysLysValSerGluPheSerLysGlnLeuGlnLys 951
 Db 3066 CAACGAGAGAAATTAAGCTGCTCAGACAGAGCTGTCTGATCTTTCTGCAACATTTGAACA 3125
 Oy 951 sAlaLeuGlnGluLysAsnAlaLeuGluAsnGluValThrCysLeuSerGluTrpLys-- 970
 Db 3126 GAAACATGCTGAGATAGTATTCTTAATGAAGAGATTAATCTTAAAGCAAGAAAGAAAGA 3185
 Oy 971 -----PheLeuProAsnGluValGluCysLeuLysAsn----- 981
 Db 3186 ACAAGTTTCATTTGAGATGTAGAGAGCTAGAAATCATTAATTAACCAACAAAGGAGCAAGAAA 3245

Oy 982 -----GlnIleSerLysAlaSerGlnGluIleMetLeuLeuGlyG1 995
 Db 3246 TGTACAGTCTGTGATTAACAGTAAAGCTTTTATTAGATGAGAGTTGTGACCATCAGAAC 3305
 Oy 995 nGluLys---GlnHISerAlaSerLysIleSerLys----- 1006
 Db 3306 CAGGGGTGCTGAAGGATCAGTTCTTAAGTAAATTAAGTTTGTGGAGAAATCAAAAAT 3365
 Oy 1007 -----GlnGluIleIleMetGlnGlnGlnSerGlnGln-- 1017
 Db 3366 AATGGTGAAGATTAAGTTCTTTTGAATAATGTGACTGTGGAGAAAGTAAACAGAGA 3425
 Oy 1018 -----IleLeuGlnLeuThrAspGluValThrHisArgLysSer-----LysValG1 1033
 Db 3426 ACAGTTGATTTTGATCACTTAACATCTGTAAACAAAGAAATGATATCAATTAAGCAACTCA 3485
 Oy 1033 nGlnThrGlnGlnGlnTrpLysGlu-----MetLysLysMetHisAspAs 1048
 Db 3486 ACCAAGTGAAATGATTAACCTTCAGAAAGAACTCAATGTAATTAATGAGAACAGATGA 3545
 Oy 1048 PheuPheGluLysTrpLleArgAsnLysSerGluAlaGlu----- 1061
 Db 3546 TTTTA-----AGCTACAGATGAGAACCCCAACCAATTTGCTCTCTGTGT 3590
 Oy 1062 -----AspLeuLeuArgGlu---MetGluAsnLeuLysGlyThr----- 1073
 Db 3591 TTATTCAACTCATGTGATCAGGTTCTGTGAATATGTGAAAAATGAAAGATTAAGCTCT 3650
 Oy 1074 -----MetGluSerValGluValLysIleAlaAspThr 1084
 Db 3651 TTGCAGTCTTAAGAAGAGCTTATTTTGTCTCAAGAGAAAGATCAAGAACTTCAGAA 3710
 Oy 1084 rLysHisGluLeuGlu---GluThrLleArg-----AspLysGluGlnLe 1098
 Db 3711 AATACACAGATTAAGTACAGTACAGTATGAAACACAGAAACAGTGTAGAAAGAAAGCC 3770
 Oy 1098 uLeuHis-----GluLysLysTrpPheP 1106
 Db 3771 TTTACATCTGCTCATTTGAGAAACTCAAAAGCACTGTCTGAAGATGTCTTATTTT 3830
 Oy 1106 eGlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLys 1126
 Db 3831 ACAGACTTATACAGCTGTC-----CTTGGAATATTTATATCTCTGCTT 3875
 Oy 1126 sLeuValGluGluAsnSerGlnAspProIleGluIleAsnArgLys----- 1141
 Db 3876 AAAATGTGAAGTAAATGAGAGAGACAAAGAAATTTCTGTGATTAATCTGAAATAGA 3935
 Oy 1142 -----HisAs 1143
 Db 3936 AGATCCAAATTAACAAAGTATATGATAGAGTTCACAGACTTTCAGAAATATATGCACAC 3995
 Oy 1143 nLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr----- 1160
 Db 3996 TCTTCTCAACAAAGTAAACAGAAATATACAAACAACTCTGTGACTTCAACAGCAGTAAG 4055
 Oy 1161 -----GluArgAsnSerLe 1165
 Db 4056 CAAGATCTGGGAGCAGACAGACAGATGATGAACCTGAATTTGAGAAAGAAACCTTCC 4115
 Oy 1165 uLysGluGlnVal-----IleAs 1171
 Db 4116 AAAGAGAGAAACAGAGTTTATTAATCAATTCATTCAGATGACCAATTTGGAAGACATTGA 4175
 Oy 1171 PheuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLysSerAspLeuGlnLys 1191
 Db 4176 TGTCAATCATTAAGCAAGATATCTTCTGCAAGATCTTGAAGAAACTTAACCTTGAGA 4235
 Oy 1191 sProLysGlnAspLeuGlu-----GluGlyG1 1200
 Db 4236 ACAAGTTCAAGATTAAGAAAGCTCATATCTCTTTGACAGCAACAGTTGAAAGAAACGTGA 4295
 Oy 1200 uValLysLeuLeuLeuGlnMetGluLeuLeuLysGlyHisLeu----- 1214

QY	1777	gaasleuasnnglunmetclumetvialmetleuglunmetglugluleuylsasnsergl	1797
Db	6306	AGAGTGGAGAAACAACTCAGTAGG--TTTATAGAGCGGAAACAAAAAAATACTGA	6362
QY	1797	nargThrValIlealaglurgsaplnleuglnaspaspIeuAargIuserValglume	1817
Db	6363	ACTAATG-----	6369
QY	1817	tSerIlegluThrGlnaspaspIeuAargIylsalactnglnalaleuglnglnIylsAs	1837
Db	6370	-----GATTTAAGACAGCAAAAACCAACGATTTGAAAAAGCATTAGA	6410
QY	1837	plyvalGlnGluLeuThrSerGlnIleSer-----Valle	1849
Db	6411	AAAAATAGAAAATTTTATAGATAGCAACCCATTGACAGACAACTGAGACAGATGTATT	6470
QY	1849	uglngluIylsIleSerleuLeuglnasnnglnmetleuThrasnValalathrValysgl	1869
Db	6471	CCAAAGGAATACAGAAACTAGAACAGACACTT-----AGGTTGTTCCTCGATTCCA	6524
QY	1869	uThrLeuSerGluArgAspaspIeuasnInserIylsglnHisleuPheSerGluIleGl	1889
Db	6525	GCCATATAGGAA-----CATCAAACTAGAGAGGTTGA	6557
QY	1889	uThrLeuSerleuSerleuIylsGluIylsGlu-----Phealale	1902
Db	6558	ACAGTTGCAAAATCCTGTAAAGAAAAAACAGACAAATGCAGTAGCGTTTGCTCTCTAA	6617
QY	1902	uglnglnalagluIylsAspIylsAlaaspAlaIalAargIylsThrIleaspIleThrGluIy	1922
Db	6618	AGACAGACCTTCAAAGGATTTATACAAAGAAAGAAAGTAAGAAATAGAAACGAGATTTCAG	6677
QY	1922	sIleSerAsnIleGluGluGlnleuLeuglnIalathrasnleuIylsGluThrleuTy	1942
Db	6678	AGTAAGAGAACTGGAGCGGCTTCTTGTGAGTCAGATACTTTTCAAAGGTAGAGA	6737
QY	1942	rgluArgGlu--SerleuIleGlnCysIylsGlnGlnleuAlaleuasnThr-----Gl	1959
Db	6738	CCGAAAAACATTGGAGCTGTAAACCTTAACCAACAAATGTCTCCATGAGATGACATTGCA	6797
QY	1959	uHisleuArgGluThrleuIylsSerIylsAspIylsalaleuGluIylsMetGluGlnGlu--	1978
Db	6798	GCGTGAAACGAGATGCCATATACAGAAAGGAAAGAGATTCAAACTTAGAAGGCATT	6857
QY	1979	-----ArgAspGluAlaIalasnIylsValIlealaleuThrGluIylsMetSerse	1995
Db	6858	AGAACACTTATAGAGAAAGAACTGAAAAATAGAAATGAGAACTTCAATATGCA	6917
QY	1995	rIleuGluGluGlnIleasnGluasnValThrThrleuIylsGluGluGluIylsGlu	2015
Db	6918	ATTAGAAATACAAAAAGGAATCTATACCCGCTACAGAAACTGTAGACGGAAAAACA	6977
QY	2015	uThrPheTyrlleuGlnArgProSerIylsGlnInserSerSerGlnmetGluGluIleuAr	2035
Db	6978	ATTATTT-----AAGATGACATGAGGAACACGGG	7007
QY	2035	gGluSerleuIylsThrIylsAspIleuGlnleuGluGluIalaglu-----	2049
Db	7008	ACTTGCCATTAAGAAATCGATGCCATGTCTACTACAGAACCAACATGCTATTGTGGAA	7067
QY	2050	-----LysGluIleSerGluAlathrasnGluIleIylsAsnleuThrIalysIlese	2067
Db	7068	ATTTGCCCAATATATACAGAAAAAGAGTAGAAATTCACCAATTAATAGCAAGTTAC	7127
QY	2067	rSerleuGluGluGluIleleuGlnAsnAlaSerIleleuasnGluIalValSerGluAr	2087
Db	7128	GAACATCCACAGCAACTT-----AAATTTACAACATATACAAAGGTATTGAAGAA--	7179
QY	2087	gGluAsnleuArgHisSerIylsGlnGlnleuValSerGluIleuGluGlnleuSerleuTh	2107
Db	7180	-----AAAAATCAACTGTATTAAGGATTTCTTGAAACCCAAATAGATG	7220

QY	2107	rleuysserlarysaprlnsalarphelaglinserylaryggluylarysrgluvalalals	2127
Db	7221	TTTTGGATGGTATCAGAAATGTGTGAAGAAATATAGAAGAAATATAGACAGCTCA	7280
QY	2127	nlvsllella-----	2130
Db	7281	TGAACTGATTTAAAAACTTCACACGGAATTGGCAAAATATTGGACAGACATCAATGA	7340
QY	2131	-----SerleuAlaglucyluileylsileuthrlysglumetaspGluPheargAs	2148
Db	7341	TGCTCATTCCTCTCAGAGAAGACAGACATTTAAAAACATCAATTGATGTGGTTATAGC	7400
QY	2148	pserlysgluserleuglnuglnuserSerHlsleuserglugluLeuCysthTrlyLy	2168
Db	7401	TGAAAAGCTGGCTTGGAGACAGCAAGTAGAAACCGCTAATGAATATCCCTCATGA	7460
QY	2168	srhr-----GluLeuglnMetLe	2174
Db	7461	AAATCTACTTAAGAAACCAATTTTAAATGAATCAAGCTACACAGGAATTTATCAGCTT	7530
QY	2174	ulysglnllyngsluAspIleasnaAsnlyleuAlagllyuVallysgluValAspG1	2194
Db	7521	AAAGAGAAACGTGAAGTGTG-----GAAAAGATTCAAGATATCCAGA	7555
QY	2194	u-----LeuLeuglnHlsleuserSerleuLyggluInLeuAspGlnI1	2209
Db	7566	GAAATAGTGTAACTGCGCTTATGATCATCTGCAGCAAAAGCAAAACCTGACATGAGT	7623
QY	2209	eglmetclu-----LeuArgAsnlylyLy-----LeuArgAsnlylyGlu	2233
Db	7626	CTTTACAGAGATGCTCTTAATCCCTAGAAATCAAGACATCTTCAAATCTTTGAAAG	7685
QY	2233	uCysgluysmetaspIleuMetgluLyggluIleuserValleuArgleu-----	2239
Db	7686	AAATGCAGAAAGTTCATATATTATTTGAAACAGAGTTGCTACAACTGACAGACAGCT	7745
QY	2240	-----MetGlnAsnG1	2243
Db	7746	TAGTCAGAAAGCACTTAGACACTTACCCAGTGTATTAACAATTAAGACATCAAGAA	7805
QY	2243	uProglnglnluAspAspArgAlaAlaglunmetaspIleuglnuserArgAsnG1	2263
Db	7806	AGGCCAGTTTAA-----ACAGAAATGCTTCAAAAGCAAGATTTGT	7844
QY	2263	ngluylleuglnluMet--GluysIleSerAlaValyTusergluGlnHlsThrLe	2282
Db	7845	AAACCTACAGAAATTACTTGAAGAAAAGCGCTGT-----	7881
QY	2282	uLeuserSerleuserSerGluLeuglnlysgluThr-----GluAlaHlslyuHlsCy	2300
Db	7882	-----GCTGTGTCAAGTCAAAATCCAACTTGAGCGAGTTCAGAAATATGTCMAATTCTG	7934
QY	2300	smetleuAsn-----	2303
Db	7935	TCAAGATATATCAAACAATTTTCATCAGAACCTGAAGACAAATATTCSGAATTTAAATCA	7994
QY	2304	-Ilelysglu-----SerleuserSerThrleuserArgSerPheGlySerLe	2319
Db	7995	ACTAAGACAAATATAGTTGGGTGCAGATATATCAGATTAACCTTGACAAATTCAGAA	8034
QY	2319	uGlnThrGlnHlsVallylsleuAsnThrGlnLeuglnThrleuLeuAsnlyuPheLyva	2339
Db	8055	AGAAAGCCAGGTGTGAATTCATCTAGTGTG-----ATTTAAGAAAAGAACAACT	8108
QY	2339	lVallyArgThrAlaAlaVallysgluAspHlsSerleuIleuysAspTrylGluLyAs	2359
Db	8109	AGAAATTCACAAAAAATATTTTAAAGAAAAGAAAGAGTGTGAAATTCAGAAACT	8168
QY	2359	pLeuAlaAlaglucInlyuArgHlsAspGluLeuArgleuGlnleuglnCylyeGlnG1	2379
Db	8169	ATTGGAGGGCAATAGAAAAACAGACAGAAAGAA-----	8205
QY	2379	nhlsGlyArglyuTrpserAspserAlaSerGlnGluLeuLy-----	2393

[illegible]

Dd	9246	GGTTTATGATAGTTCATCTCATATGATGAGCTTTCAGACTGGCAGGTGAACCTACTGCT	9305
Oy	2575	allleuglinasprysleu	2582
Dd	9306	TGCCCTTCACAAAGTTTTCTTGAAGAAGCTAGTGTTTACTAGCAGCATTTTCGACGCA	9365
Oy	2583	-----SeratgasaalaglualagluleuAsnalAmetGlnVallyle	2597
Dd	9366	GCTGACACCTCTAGACTCTACAGATGCGAGTGGTTTACTTACCTAACCTGTTGGAAACAGAGAT	9425
Oy	2597	uthrltyslsglinaspasleuglnalaametlys	2609
Dd	9426	KCAAAACAGAGGTGTGTAATATCAACGACTATGGAATGCTCCAGAAAACAGATAGANG	9485
Oy	2610	-----GulilegiasleuGlnlyMetValaIalalyalavalproty	2625
Dd	9486	GAGTTTGTATTCTGAAATCTGACGCACTGTCATGCAAAATGAATGTAAGGAAATTTACTCT	9545
Oy	2625	rltysglugluleap	Asnleulstth 2634
Dd	9546	GAAAGAGAGAACAGAGAGTGAAGAAACCAAGCAAGACTCTTGAAATATATATATACAGA	9605
Oy	2634	rltys-----ValvalyslleglumetGlnlytleystysrserlysal	2649
Dd	9606	GAAAGCAGTCTCAATGCTGAGATGCAAGTGAAGTCAAGCAGATGTGAA--GACAGAGC	9662
Oy	2649	athrlasplnglnlualatytleulysSerCysleuglnasprysglugluleulr	2669
Dd	9663	AACGGAA-----CTGCAGAGACAGCTGAGTTCTGAGAAATGGTGTGTC	9707
Oy	2669	gatlleulysglugluleulatyrtalaglnalalaspaslnsprthrlthralCysvalpr	2689
Dd	9708	TGAACGTGAAGAGTCACTGCACAACTAAATGGAACTGTAACAAACACCTCAAGGCACA	9767
Oy	2689	olysaspryglulysalaserthpheprovalthrCysgllygllyserglyleval	2709
Dd	9768	GCAATAACACCTTAA--	9783
Oy	2709	lglnserthralametleuValleuglnserglulysalalaleuglnulrGluileuse	2729
Dd	9784	-----GAATTGAGGCTTTTCAGGTTGGAAAGTTAA	9812
Oy	2729	rhlttyrlylysalstytHshltisleuseratrgthmetserSerSerGlnaaparg--	2748
Dd	9813	AGATTAACAGACAGTGAATCATTTGGTTATGACACATTAGCAAGTGAACAGAAATAATC	9872
Oy	2749	-----lystysThrlysalalysSerapralahlsSerSerH	2761
Dd	9873	AAGAGACTCCAGTGGGCTTTGGAGAAAGAAACGA--	9912
Oy	2761	stthglyserSerthrlasrglyserProhlystysThrGlnulthrtYarghlsGly--Pr	2780
Dd	9913	-TTGGGAGCGCAGTGAAGACGGGATAAGAAAGAACTGAGAGATTCGAACTTTCACCTGA	9971
Oy	2780	ovalthrProglulnrgserglumetProserleuHlsleuglyserProlyglySerGcl	2800
Dd	9972	GAGTCAGAAACAAAGAAATCTTCAGCTAAATCTACTTTTGGAAACAACAAACAATCTCT	10031
Oy	2800	uSerSerthrltysrlyvalalaserProasnatysSerGlnllytysSerGlnleuValme	2820
Dd	10032	GAAACAAATCCAGCAAAAATATGACATACACAGAAATGATATGATGCCCACTTGCAGA	10091
Oy	2820	tSerProgllysthrGlymetHlslyshlsleuSerProserlysalGlyleuHl	2840
Dd	10092	AGAACAAAGTGAACAACTTAGAGCTTTCAGTACTTGTGATCTGCAAAAAGTTGCAATTCG	10151
Oy	2840	sltysltsrghalaleuserProasnatysSerGlnumetProthrlGlnlsvalIlaserPr	2860
Dd	10152	CGAAATGAGTAGTACCTAGATAGGGAACGGGAA--	10185
Oy	2860	oGlylsthrlglyleuHlslstysAsnleulthrlGlnSer--thrlleupheasplnleuS	2880
Dd	10186	-----TTGCACGCGACAGCTGCAGAGCAGTGAATGATCTGACAGCTCTCGGCC	10232

OY 2880 eProCysLysGlnGln 2885
 DB 10233 ACCCTTGCCCTCAGAG 10249

RESULT 12
 US-10-082-830-134
 ; Sequence 134, Application US/10082830
 ; Publication No. US20030077604A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sun, Yonmgling
 ; APPLICANT: Recipon, Hevea
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Turner, Leah
 ; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
 ; FILE OF INVENTION: Genes and Proteins
 ; FILE REFERENCE: DEX-0249
 ; CURRENT APPLICATION NUMBER: US/10/082,830
 ; CURRENT FILING DATE: 2001-10-29
 ; PRIOR APPLICATION NUMBER: 60/243,802
 ; PRIOR FILING DATE: 2000-10-27
 ; NUMBER OF SEQ ID NOS: 282
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 134
 ; LENGTH: 11677
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-082-830-134

Alignment Scores:
 Pred. No.: 1,12e-54 Length: 11677
 Score: 998.00 Matches: 599
 Percent Similarity: 39.048 Conserves: 521
 Best Local Similarity: 20.888 Mismatches: 1011
 Query Match: 6.768 Indels: 739
 DB: Gaps: 114

US-09-150-867-1 (1-2954) x US-10-082-830-134 (1-11677)

OY 218 AlaGluGlyValArgLeuLysGluGlyCysAsnIleAsnArgSerLeuPheIleuGly 267
 DB 4421 GGTGAGGCGC-----CTTAGCAGCGCGCTGCCAGCTCAATCGCGAGCACTAGCGCGG 4474
 OY 268 GlnValIleLysLysLeuSerAspGlyGlnAlaGly----- 279
 DB 4475 CGT-----CTGCGGAGCCGAGCGCTGGGACCTGAGGCTTACCACTGGTCCCTCATGAG 4531
 OY 280 -----GlyPheIleAsnTyrArgAspSerLysLeuThrArgIleLeuGlnAsn 295
 DB 4532 ACAAGAGCGCTGGGTGTAACAAACATGAAGCCCAAGTCACTCAGCTGTACTGGAGAG 4591
 OY 296 SerLeuGlyGlyAsnAlaLysThrValIleIleCysThrIleThrProValSerPheAsp 315
 DB 4592 CAGGTG-----CTGCACACTACAG 4609
 OY 316 GlnThrLeuSerThrLeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHis 335
 DB 4610 CAGCAGATGGCGAAGAAATCAGCAGCCTCTCGCGGAG-----CTGAAGAACTCCAGAGAG 4666
 OY 336 ValAsnGlu-----ValLeuAspAspGluAlaLeuLeuLysArgTyrArg 350
 DB 4667 GCCCGCAGAGCAGCAACACCTTGTGAGAGAGCTGAGGCGCAAGGTGCTGCAGTACCGA 4726
 OY 351 LysGluIleLeuAspLeuLysLysGlnLeu----- 360
 DB 4727 AGCTGTGCCAAGAGCTGGAGAGCGCTTGAAGCCACTGAGAGCAATCCCGCAGAG 4786
 OY 361 ---GluAsnLeuGluSerSerSer-----GluThrLysAlaGlnAlaMetAlaLysGlu 377
 DB 4787 TGGGAAATATGTGAGAGAGCAACCTCGATGAGCTGTGCTCGGATGTGAGAGAGAGCA 4846
 OY 378 GlnHisThrGlnLeuLeuAlaGluIleLys-----GlnLeuHisLysGluArgGlu 394

DB 4847 CAGAGCTGTGAGAGCTTACACAGGTGAACACCCAGCTTCGACTGCATGTGAAAAAGCT 4906
 OY 395 AspArgIleThrPheIleThrAsnIleValAlaSerSerGlnLysSerGlnAsp 414
 DB 4907 GAC-----GTGTGAATATAAGCCCTTAGAGAAATGATGTGMA----- 4942
 OY 415 GlnArgValLysArgLysArgValAlaThrThrPalaProGlyLysIleGlnAsnSerLeu 434
 DB 4943 -----AACTGACAGTGAAGTGAAGCGGCGCG----- 4972
 OY 435 HisAlaSerGlyValSerAspPheAspMetLeuSerArgLeuProLysAsnPheSerLys 454
 DB 4973 -----GATGAGCTTAATAGAGAGAGAGAGCCAGTGCAGATG 5008
 OY 455 LysAlaLysPheSerAspMetProSer-----PheProGluIleAsp 468
 DB 5009 GAGCAGAGAGTGGAGTCTGTTTCTGCTTGTCTGCTGATCAGAGATCTGATGAGCTAATA 5068
 OY 469 AspSerValCysThrGluPheSerAspPheAspAspAlaLeuSerMetLeuAspSerAsn 488
 DB 5069 GCTGA-GCATGTGAGGCTTTCAGAGCTCTGTGACCTGTGTGTGCTGCTGCTGCTGCTG 5127
 OY 489 GlyIle-AspAlaGluThrPAsnLeuAlaSerLysValThrHisArgGluLysThrSerLe 508
 DB 5128 AGCAGAGCTCGGAGAACCCACAGGATCTGGAAGATGAATGAGTGGCGGAGCCGCCAGCT 5187
 OY 508 HisGlnSerMetIleAspPheGlyGlnIleSerAspSerValGlnPheHisAspSerLe 528
 DB 5188 GCTGCTGCTACTAGCCAGACCCAGCCAGCTGTGAGAG-----GAAGCCCATGMAAGAG 5241
 OY 528 rLysGlu--AsnGlnLeuGlnTyrLeuProLysAspSerGlyAspMetAlaGluLysAr 547
 DB 5242 CAGGAGTTAATACAGCTG-----NAGACTCAAGGGATCTG-----GA 5280
 OY 547 gLysAlaSerPheGluLysGluIleThrSerLeu-----GlnGlnGlnLe 562
 DB 5281 GAAGCTGAACCTCAGAGCGGGGTGACGAGCTCTGCTGTTGACCCAGTCTAGAA 5340
 OY 562 uGlnSerLysGluGluGluLys-----LysGluLeuValGlnSerPheG1 577
 DB 5341 GCMAATGAAGATTATGAAGAAATGATTAAGCTCTGAGAGAGACGATGATCTGTGA 5400
 OY 577 uLeuLysIleAlaGluLeuGluGluGlnLeuSerValLysAlaLysAsnLeuGluMetVa 597
 DB 5401 GACAATTCACACAGATTTATGGAACAT-----GAAGCATCTCTTGA 5442
 OY 597 ThrAsnSerArgGluHisSerIleAsn--AlaGluValGlnThrAspValGluLysG1 616
 DB 5443 TAGGAATGGCGAAGAGAGAGATGTGTTTACAGCAGGTGATCAAGATATTAACCCAGGT 5502
 OY 616 uValValArgLysGluMetSerValLeuGluLysPheSerGlyTyrAsn----- 631
 DB 5503 CATGTGTGAAGAGGGGAGCAATATACCCAGAGCTCTGTGATAGAACTTTTGAAT 5562
 OY 632 -----AlaSerAsnSerAspLeuGlnAspSer----- 640
 DB 5563 GGACTCTACTATCTTCTCCAGTTGATTAACAGATGACAGACAAAGCTTTACTCTGTGT 5622
 OY 641 ----SerValAspGlyLysArg----- 646
 DB 5623 GCGTTACAGCTGACACCGAGAGCCAGCTGTGACAGACTTAAGGACAGACCTTGACAG 5682
 OY 647 -----LeuSerSerSerHisAspGluCysIleGlnHisArgLys 659
 DB 5683 CTGTCAAGAGGCTGTGAACCTTGTGCACAGCAGCATGATGATGAGAGAGAGGCGCA 5742
 OY 659 smetLeuGluGlnLysIle----- 665
 DB 5743 ACCCTTGAACAGCGGCTGCAGAACTCACTGGGAGCGGACACTCTGCGAGGCGACAGC 5802
 OY 666 -ValAspLeuGluGluPheIleGlnAsnLeuAsnLysLysSerGlu----- 680

Db 5803 TGTGACCTCCAGAGAGGTGACTCTCTCAGCAAGAGGAGAGAGCTGTCTGAGAGGC 5862
QY 680 ----- 680
Db 5863 CAGGAGAGAGCTCGCGCAGAGCTGGAGTGTCTAGACAGAGAGAGCTGGCGCTGGAG 5922
QY 681 ----- 692
Db 5923 GGTAAATGTGAGCTTCAGCTGCAGGGGAGACTCTGCCAGGCGCAAGAGAGAGAGAGCA 5982
QY 692 TGLUSETLLEGLNLEUCYS-----GluAlaIleMet----- 702
Db 5983 GGAGGAGCTGCACCTGCTGTCCGGAGAGGAGAGAGCTTTCAGAGAGATGCTGTATGGCCCT 6042
QY 703 -----AlaGluLysAlaAsnAlaLeuGluLysLeuAlaLeuMetArgAspAsnPhaAsp 721
Db 6043 GGAAAGCCAAACAGTCACAAATCTACTAGTAACATGATCATCTTGGGAAAGCCCTGGAGTGC 6102
QY 721 nilelleuGluLysnGluThrLeuLysArgGlu-----IleAl 734
Db 6103 AAGTCACCTGGAGAGGGGAGTACTAGGAGCAAGAGCAAGAGAGAGTGCAGCGAGCTGGC 6162
QY 734 aspleuGluArgSerLeuLysGlu-----AsnGluLysThrAsnGluPheGlu 751
Db 6163 TAGGGCAGAGCAAGTCATATTGCAGAGCTGTGAGTTCTCAAAACCCCTGAAGAGAGAGT 6222
QY 751 eleuGlu-----LysGluThrGlnLysGluHisGluAlaGluLeuIleH 766
Db 6223 AGCTATCTTCGGGCTGCAGCTGCATGACCTTAAATGAGAGCTTGGGCTTGA 6282
QY 766 sGluIleGlySerLeuLysLysLeuValGluAsnAlaGluMetArgAsnAlaLeuGlu 786
Db 6283 TAAAGTTGGCTGAACAGCAGCTTCTCCAG-----TTAGAGAGAGAGAGAGAGCTGTGTG 6339
QY 786 uGluAspleuGluThrLysLysLeuLysGluGluLysGluIleGluLeuAlaGluLeu 806
Db 6340 CAGCAGAAATGGAGGCGGAG 6399
QY 806 uArgLysArgAlaAspAsnLeuGlnLysLysValArgAsnPhaAspleuSerValSer 826
Db 6400 AGAGAGAGAG----- 6409
QY 826 tGlyAspSerGluLysLeuGluGluLysLeuPheGluLeuLysGluSerLeuSerAspAl 846
Db 6410 -----AGGAG 6462
QY 846 agluAlaValThrArgAspAlaGlnLysGluLysSerPheLeuArgSerGluAsnLeuGlu 866
Db 6463 TGGAGAGAGCTGGGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6522
QY 866 uLeuLysGluLysMetGluAspThrSerAsnTrpTyrAsnGlnLysGluLysAlaAlaLase 886
Db 6523 AATTTCAAAAGAACTAAGTACAGT-----CACAGAGAGAGAGAGAGAGAGAGAGAGAG 6573
QY 886 rLeuPheGluLysGluLeuGluLysSerAsnTrpTyrLysLysMetGluAlaAspleu 906
Db 6574 TCAGCTGAG-----CAGCTTACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6630
QY 906 uGlnLysGluLeuGlnSerAlaPheAsnGluLysLeuAsnTrpTyrLysGluLysAlaGlu 926
Db 6631 AGTCCAG 6690
QY 926 yLysValProArgAspleuLeuSerArgValGluLeuGluLysLysValSerGluPhe 946
Db 6691 C-----GTGAG 6747
QY 946 rLysGlnLeuGluLysAlaLeuGluGlu----- 955
Db 6748 GGAGCTACTGAG 6807
QY 956 -LysAsnAlaLeuGluLysnGluValThrCysLeuSerGluTyrLys-----PheLeuProAs 974
Db 6808 CAAAGGAG 6867

QY 974 nGluValGluLysLeuLysAsnGlnLysSerLysAlaSerGluGluLysLeuLeuLys 994
Db 6868 GGAAGTGAAGTGCCTGAG 6906
QY 994 sGlnGluLysGlnHisSerAlaSerLleIleSerLysGlnGluLysLeuIleMetGlnGlu 1014
Db 6907 GAGTCAAGCAG 6960
QY 1014 nSerGluGlnLysLeuGlnLysLeuThrAspGluValThrHisThrGlnSerLysValGln 1034
Db 6961 GCAG 7020
QY 1034 nThrGluGlnGlnLysLysLeuLysLysMetLysLysMetHisAspAspleuPheGlu 1054
Db 7021 GCTCCGAG 7060
QY 1054 eArgAsnLysSerGluAlaGluAspleuLeuArgGluMetGluLysnLeuLysGlyThr 1074
Db 7061 -----CAG 7101
QY 1074 tGluSerValGluValLysLysLysAlaAspThrLysHisGluLeuGluGluLysThrArg 1094
Db 7102 AATGAG 7161
QY 1094 rLysGluGlnLysLeuHisGluLysLysThrPhePheGlnAlaMetGluThrLysPhe 1114
Db 7162 GAG 7210
QY 1114 oIleThrProLeuSerAspSerLeuProSerLysLeuValGluLysnSerGlnAs 1134
Db 7210 ----- 7210
QY 1134 pProIleGluLysAsnAspTyrHisAsnLeuIleAlaLeuAlaThrGluArgAsnAl 1154
Db 7210 ----- 7210
QY 1154 eMetValCysLeuGluLysThrGluArgAsnSerLeuLysGluGluValLysAspleuAsn 1174
Db 7211 -----CTGAG 7248
QY 1174 rGlnLeuGlnSerLeuGlnAlaGlnSerLleGluLysSerAspleuGlnLysProLysGlu 1194
Db 7249 ACTGCTCAG 7308
QY 1194 nAspleuGluGluLysLysValLysLeuLeuGluMetGlu-----LeuLeuLys 1211
Db 7309 GGAGATG-----AAAGTCAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7359
QY 1211 sGlyHisLeuThrAspSerGlnLeuSerLleGluLysLeuGlnLeuGluLysnLeuGluVal 1231
Db 7360 GAGCCAGCTCCAG 7411
QY 1231 lThrGluLysLeuGlnLysLeuGlnGluGluLysAsnLysLysLysLysLysLysLys 1251
Db 7412 -AGAGATGACCTTCTGCTCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7470
QY 1251 uLeuGlnLysLeuThrAspSerLysLysAlaGlnLysAspSerLeuLysGlnLysLeu 1271
Db 7471 CCTGAG 7524
QY 1271 rGluAsnLleGluGlnSerLleGluThrGlnAspGluLeuArgAlaAlaGlnGluGluLeu 1291
Db 7525 CCAG 7581
QY 1291 uArgGluGlnLysGlnLeuValAspSerPheArgGlnGlnLeuLeuLysPysSerVal 1311
Db 7582 CCGAATTCAGAGAGAGAGAG-----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7623
QY 1311 yLysSerProAsnHisAspAlaValAlaAsnGlnLysLysValSerLeuGluGluVal 1331
Db 7624 TCTCTCAGTGAAGAGAGAGAG-----CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7671

Db 9550 CACTGAGGACGCTGATCAGGAAACGGCAGAGGAAGGAGGCGC----- 9595
 QY 2017 eTyrLeuGlnArgProSerLysGlnInSerSerGlnMetGlnLeuArgGlu 2037
 Db 9596 -----CCAGTAAAGACACAGCGCGGAGC-----CTAGACACATGAAAGCTGAT 9639
 QY 2037 rLeuLysThrLysAspLeuGlnLeuGlnLysGlnLysSerGlnLysThrAs 2057
 Db 9640 CCGTGGTGTAAAGAGAAAGAGGTGAAATGATGAGCAGAGCATATCCATGAACTCCAGGA 9699
 QY 2057 nGluLeuLysAsnLeuThrAlaLysIleSerSerLeuGlnGlnLysLeuGlnAsnAl 2077
 Db 9700 GGTCAAAAGCCAGCTGGACACAGACAGCTCCAGGCGCTGCACAGAAAGTAGTGACACCG 9759
 QY 2077 aSerIleLeuAsnGlnAlaValSerGlu-----ArgGluAsnLeuArgHisSe 2093
 Db 9760 CCTCTCTCTGTCCTCCAGCAGACAGAAATAGTGTCTCTGCACAGCAACTCCAGGAAC 9819
 QY 2093 rLysGlnGlnLeuValSerGlnLeuGlnGlnLeuSerLeuThrLysSerArgAspH1 2113
 Db 9820 CAGGGAACAA-----GGGAGCTGAGAGGACGATCATT----- 9853
 QY 2113 sAlaPheAlaGlnSerLysArgGluLysAspGluAlaValAsnLysIleAlaSerLeuAl 2133
 Db 9854 -----CAGAGTCAACTGTGATGAGCCCAAGACGCCCTAGCCACAGAGGA 9897
 QY 2133 aGluGlnLysIleLeuThrLysGlnMetAspGluPheArgAspSerLysGlnSerLe 2153
 Db 9898 CCAGAGACGTGAGGCTGTCAGACAGAAACAGACAGAGCCGAGGAGAGAGAGGT 9957
 QY 2153 uGlnGlnInSerSerHisLeuSerGlnGlnLeu-----CysThrTyrLysThr 2169
 Db 9958 GAAGAAAGGACAGACGCCCTCCAGGAGCTCTGGAGCAAGCCCATGATGACTGAAGA 10017
 QY 2169 r-----GluLeuGlnMetLeuLysGlnGlnLysGlnLysAsnAlaSerLeuAl 2186
 Db 10018 GCGTATGAGAGCTTTCAGACACACAGAAACAGACAGAGCCGAGGAGAAAGCTGCG 10077
 QY 2186 aGluLysValLysGlnValAspGlnLeuGlnHisLeuSerSerLysGlnGlnLe 2206
 Db 10078 AGTGAAGGAGCGCGGCTC-----CAGGCGCTGAGGAGGAGTCT 10116
 QY 2206 uAspGlnIleGlnMetGlnLeuArgAsnGlnLysLeuArgAsnTyrGlnLeuGlnLys 2226
 Db 10117 GAGAGACTAAGGGCTGAGTCTCG----- 10141
 QY 2226 sMetAspIleMetGlnLysGlnLysSerValLeuArgLeuMetGlnAsn-----GluPr 2244
 Db 10142 -----GAACAGAGAAAGCTCTGTGGCTCCAGACAGCAGTGTGAGCA 10188
 QY 2244 oGlnGlnGlu-----GluAspAspValAlaGlnArgMetAs 2256
 Db 10189 GGCACAGAGCATGAGTGGAGACACAGGCGCTTCGACAGACGTCGGTGCAGGCCAGGC 10248
 QY 2256 rIleLeuGlnSerArgAsnGlnGlnIleGlnLysLeuMetGlnLysIleSerAlaVal 2276
 Db 10249 AGTCTCAAGAGACGGAGCAGCAGTGTGAGCAAGTCTGCGGCAAA----- 10294
 QY 2276 rSerGlnGlnHisThrLeuLeuSerSerLeuSerGlnGlnLysGlnLysGlnThrGlnAl 2296
 Db 10295 -----AGTCAAGTCTCTCCCGGATCAGAGAGGAGGCTGCGCG 10329
 QY 2296 aHisLysHisCysMetLeuAsnIleLysGlnSerLeuSerSerThrLeuSerArgSerPh 2316
 Db 10330 GCGCCCG-----GCTGAAGCTCTGCAGAGGCGCTTCGAGAGGTCA 10371
 QY 2316 eGlySerLeuGlnThrGlnHisValLysLeuAsnThrGlnLeuGlnThrLeuLeuAsn 2336
 Db 10372 TGCTGCGCTGACGAGGAAG-----CAGATCTCTC----- 10405
 QY 2336 sPheLysValValLysArgThrAlaAlaValLysGlnAspHisSerLeuLysAspTyr 2356
 Db 10406 -----GAGCAGGACAGAAATTTAGCGCGCACTCT 10431

QY 2356 rGluLysAspLeuAlaAlaGlnGlnLysArgHisAspLeuLeuArgLeuGlnGlnC 2376
 Db 10432 GAGGCCACAGCTGACACC-----CTGCAAAGCTCTCCCTGATGC 10470
 QY 2376 sLeuGlnGlnHisGlnArgLysTrpSerAspSerAlaSerGlnGlnLeuLysPheCysG 2396
 Db 10471 CTGCCAGGCACACAGTGGCAG-----CTGAGAGAGGCTCTGAGAGATCAAGA 10518
 QY 2396 uIleGluPheLeuAsnGlnLeuLeuPheLysLysAlaAsnIleIleGlnSerValGlnAs 2416
 Db 10519 AGGTGAGATTCACAGACCCAGATCTC-----CGATACACAGA 10554
 QY 2416 pAspPheSerGlnValAlaValPheLeuAsnGlnValGlySerThrLeuGlnGlnGln 2436
 Db 10555 GATGTGCAGACAGCTGCACGACGACCTTGCCAG-----AGGATGAAGAGCT 10602
 QY 2436 uGlnHisLysLysGlnPheMetGlnTrpLeuGlnGlnPheGlnLysPheuHisValAspAl 2456
 Db 10603 GAGACATCAGACAGAACGAGGAGCAGCTGTGAGAG-----TCTCGGC 10647
 QY 2456 aLysLysLeuSerGlnGlnMetGlnGlnLysAsnArgArgIleAlaSerThrIleGlnLe 2476
 Db 10648 CCAGAGGCTCCAAAGAGATATGATTCAGAGAACAGAAATCTGGCGCAGAGAGAA-- 10705
 QY 2476 uLeuThrLysArgLeuLysAlaValAlaGlnSerLysIleGlnArgGlnIleThrValTy 2496
 Db 10706 -----GAGAGAGAGATTAAGGGCTTCACTACAGT-----CTAAGGAGCTACAGCTGAC 10755
 QY 2496 rLeuAsnGlnPheGlnAlaLysLeuGlnGlnLysLysGlnGlnAsnLysGlnLeuMetAr 2516
 Db 10756 TCTAGCCCAAAAGAGACAGGAGATTCTGAGCTGAGGAGAACCCAGCAAGAAACAACCT 10815
 QY 2516 gArgMetGlnHisGlnProSerAlaSerValMetGlnGlnGlnLysAsnAlaArgLeuLe 2536
 Db 10816 GGAAGCCTTACCCACAGCCACAAAACTCCCAATGAGAGAACAACTTAATAACTGA 10875
 QY 2536 uGlyTleLeuLysThrValGlnAspGlnSerLysLysLeuGlnSerArgIleLysMetLe 2556
 Db 10876 TTCTTTAGAGCCCGAGCTGAGCGGAGCTGAGACGGCTACAGACACCTGAGACAGAC 10935
 QY 2556 uGluAsnGln-----LeuAsnLeuValLys 2564
 Db 10936 AGAAGCCAGAGAGATTGAGTGGAGGAGAAAGCCACAGCTGAGCACTTCCCTAGCGCA 10995
 QY 2564 sAspAspAlaMetHisLysGly-----GluLysValAlaIleLeuGlnAspLysLeuLe 2582
 Db 10996 GACCAAGGCCAGTGTGACAGTCTGCAGAGGTAGCCATGTCTTACACAGCTCTGTCT 11055
 QY 2582 uSerArgAsnAlaGlnAlaGlnLeuAsnAlaMetGlnValLysLeuThrLysLysGlnAs 2602
 Db 11056 GGAAGCGGAGCTCAGAACAGCA-----AGCTGCAGGA 11088
 QY 2602 pAsnLeuGlnAlaAlaMetLysGlnIleGlnLysLeuGlnLysMetValAlaLysGlnAl 2622
 Db 11089 TGAAGTGAAGCTCAGACAGAGGCTCTGAGAG--GAGCGCTACACAGCCAGAGTGC 11145
 QY 2622 aValProTyrLysGlu-----GluLeuAs 2630
 Db 11146 AACCAAGACAGCAAGAACTGGGCTCCAGAGGCGAGCAGGGGTCTCAGCTGGAGAGGTTC 11205
 QY 2630 pAsnLeuLysThrLysValValLysIleGlnMetGlnLysIleLysTyrSerLysAlaThr 2650
 Db 11206 AGGAGTGGAGCTGAGCTAGTCTGATGGAATGGAGAAAGCACTCATG----- 11254
 QY 2650 rAspGlnGlnIleAlaTyrLeuLysSerCysLeuLysAspLysGlnGlnGlnArgAr 2670
 Db 11255 -----AGACAAAGGCTTGAACACCTGCACACAGCAAGCTGCGCG 11292
 QY 2670 gLeuLysGlnGlnLeuArgArgAlaGlnAlaAspAsnAspThrThrValCysValProLy 2690
 Db 11293 GCTGGAAGATTGACAGAGACAGGCTGACGCGCCACAT----- 11329

QY 2690 sasPTyrglnLysAlaSerThrPheProValThrCysglYglYSerGlyIleValGI 2710
 DB 11330 ----GTCCAGCTGCGGAGTACCTTGGAG-----CAGGATGGAGGAGCAGAAAGAA 11376
 QY 2710 nserThrAlaMetIleVal-----LeuGlnSerGluLysAlaAlaLeuGluArgGlu 2728
 DB 11377 CTCAGATGCCAAGTGTGTGGCTGAACTGCAGAAAGAGGTGCTCTCCACACTGAGCT 11436
 QY 2728 user---HisTyrlsLysLysTyrlsHisLeuSerArgThrMetSerSerSerGluAs 2747
 DB 11437 GACTTTGGAGCGGAGCAGAAAGCAGACGACTACATCACCCGCTGACGACAGACCCGCTGA 11496
 QY 2747 pArgLysLysThrLysAlaLysSerAspAlaHisSerSerHisThrGlySerSerHisAr 2767
 DB 11497 G-----CTAGCAGGCGCTGCACCA 11514
 QY 2767 gglYSerProHisLysThrGluThrTyrlarGHisglYProValThrProGluArgSerGI 2787
 DB 11515 CAGCCTTCACACTCTTCTGCGCGCCGAGCCGCT-----GA 11556
 QY 2787 umetProSerLeuHisLeuGlySerProLysLysSerGluSerSerThrLysArgValVa 2807
 DB 11557 GGGCACTGCTG-----GAGGCAAGAACCCGCGAGCTG-- 11590
 QY 2807 lserProAsnArgSerGluIleTyrlSerGlnLeuValMetSerProGlyLysThrGlyMe 2827
 DB 11591 -----GATGAGTCCCTGACTCAAGTCTGACATCCCGAGGCGAGTCTGCT 11637
 QY 2827 thlslYshlslleuSerProSer 2835
 DB 11638 ACACCC-----AGCCCGAGC 11653

RESULT 13

US-10-098-841-40

Sequence 40, Application US/10098841

Publication No. US20020197679A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Xu, Chongjun

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhao, Qiang A.

APPLICANT: Ren, Relyan

APPLICANT: Chen, Rui-hong

APPLICANT: Wang, Duntui

APPLICANT: Wang, Zhiwei

APPLICANT: Wehrman, Tom

APPLICANT: Zhang, Jie

APPLICANT: Qian, Xiaohong B.

APPLICANT: Drmanac, Radjoje T.

TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and

FILE REFERENCE: 784CIP2

CURRENT APPLICATION NUMBER: US/10/098,841

PRIOR FILING DATE: 2002-03-13

PRIOR APPLICATION NUMBER: 09/598,042

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 331

SOFTWARE: PC_FGenes Version 1.0

SEQ ID NO 40

LENGTH: 6386

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (150)..(5492)

US-10-098-841-40

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-150-867-1 (1-2954) x US-10-098-841-40 (1-6386)

5 AspAlaValLysValCysValArgValArgProLeuIleGlnArgGluGlnGly----- 22

DB 318 GATTATCTCAGGTTGTGTCTGCAATTAAGCATTATACACGTCAGAAAAAGAACTTGAG 377

QY 23 -----AspGlnAlaAsnLeuGlnInTrpLysAlaGlyAsnAsn 34

DB 378 TCTGAGGCTGTGTGATATATCTGATTCACACACTGTGTCTGAAAGAGCCTCAATGC 437

QY 35 ThrIleSerGlnValAspGlyThrLysSer-----PheAsnPheAsp 48

DB 438 ATCTTGCTGCTGTTAAGTCAAAAAAGCTCAGGCAATGCGCACAGAAATTCAGTTTTC 497

QY 49 ArgValPheAsnSerHisGluSerThrSerGlnIleTyrlGlnLysAlaValProIle 68

DB 498 AAGGTTTGGCCAGCAACTACACAGAAAGAAATTCCTTCAGGTTGCATTATGCACCA 557

QY 69 lIeArgSerAlaLeuGlnGlyTyrlAsnGlyThrIlePheAlaTyrlGlyInTrpSerSer 88

DB 558 GTAAAGACCTCTTGAAGAGCAGAGCTGCTGATTTTACTTACGCGGCAACCAATTC 617

QY 89 GlyLysThrTyrlrMetMetGlyThrProAsnSerLeuGlyIleIleProGlnAlaIle 108

DB 618 GGAATAACATATACATTTCAAGGACAGAAAGAAATTTGGCATTCGCTCGAAGCTT 677

QY 109 GlnGluValPheLysIleIleGln-----GluIleProAsnArg 121

DB 678 AATGTTATTTGATATGATGTCTCAAGAAAGCTATACAAAGATGAACCTTAACCAAT 737

QY 117 ----- 126

DB 738 AGATCCAGAGAAATCTTAAGTTATCATCAGAACAGAAAGAAAGAAATTTGCTAGCA 797

QY 122 GluPheLeuLeuArg----- 126

DB 798 AGTGCATGCTTGCGAAATTAAGAGGTTACTGTGCATTAAGATGATGATCTCTT 857

QY 126 ----- 126

DB 858 TATGAAAGTTTACTAAGCTTTGAATATCTCAGAGTTTGAAGATTCATTAAGATTA 917

QY 127 -----ValSerTyrlMet 130

DB 918 GAACAAGCCAACTTGAATATGCTTAATATTAATTAATTTCTGTGTGCTTCTTCTTT 977

QY 131 GluIleTyrlAsnGlnThrValLysAspLeu-----CysAspAspArgArg 146

DB 978 GAAATTTACAAATGATATATTTATGACTTATTTTCTTCCTGTATCATCTTAATCCAA 1037

QY 147 LysLysProLeuGlnLysArgGluAspPheAsnArgAsnValTyrlValAlaAspLeuThr 166

DB 1038 AGAAAGATGCTGCGCTTCCCAAGCGTAAGGCGTATTTCTTTTAAAGATCTACAA 1097

QY 167 GluGluLeuValMetValProGluHisValIleGlnTrpIleLysLysGlyGluLysAsn 186

DB 1098 TGGATTCAAGTATCTGATTCACAAAGAGCTATAGACTTTTAAACTAGAAATTAAG 1157

QY 187 ArgHisTyrlGlyLysThrLysMetAsnAspHisSerSerArgSerHisThrIlePheArg 206

DB 1158 CAGAGTGTGCTTCAAAATTAATTAATGCTTCAGTAGAAGTACAGCATATTACCT 1217

QY 207 Met---IleValGluSerArgAspArgAsnAspProThrAsnSerGluAsnGly 225

Db 1218 GTTAAATATTACAGATTGAAGAT-----TCTGAATGTCCTGT 1256
 OY ::::: |||
 OY 226 AAlaValMetValSerHisLeuAsnLeuValAspLeuHlaGlySerGluAlaSerGln 245
 Db 1257 GTAATTCAGACAGATGATTTATCTTTATGTCATCTTCTGTCGTGAGCAAGACTGAGAC 1316
 OY ::::: |||
 OY 246 ThrGlyAlaGluGlyValArgLeuLysGluLysCysAsnIleAsnArgSerLeuPheIle 265
 Db 1317 ACACAGAAATGAGGTGAAAGGTGAGAGACTGGAAATATCAACACTCTTTATTGACT 1376
 OY ::::: |||
 OY 266 LeuGlyGlnValIleLysLysLeuSerAspGlyGlnAlaGlyPhe-----Ile 282
 Db 1377 CTGGGAAAGTGTATTACGCTTGTGAAGATAGTGAAGATCAAAAGTCAACACCATGTG 1436
 OY ::::: |||
 OY 283 AsnTyrArgAspSerLysLeuThrArgIleLeuGlnAsnSerLeuGlyAsnAlaLys 302
 Db 1437 CCTTCCCGGAAAGTAACTGACTACTATTCTCAAACTTTTAAATGTAAGGGAAA 1496
 OY ::::: |||
 OY 303 ThrValIleIleCysThrIleThrPro-----ValSerPheAspGluThrLeuSerThr 320
 Db 1497 ATTGTGATGATGTCAATATTCAGCCAAATGTTATTAGCCTATGATGAAGAACTCAATGTA 1556
 OY ::::: |||
 OY 321 LeuGlnPheAlaSerThrAlaLysHisValArgAsnThrProHisValAsnGluValLeu 340
 Db 1557 TTGAAGTCTCCGCGCATTCACAAAGAGTTGTGTCCACAGACACTTTAAATTCCTCTGAA 1616
 OY ::::: |||
 OY 341 Asp-----AspGluAlaLeuLysArgTyrArg 350
 Db 1617 GAAAGATTTATTGGACCTGTCAAAATCTCTCAAGATGTATCTACATGACAGTAATTCAAAC 1676
 OY ::::: |||
 OY 351 LysGluIleLeuAspLeuLysLysGlnLeuGlnAsnLeuGluSerSerGlnThrLys 370
 Db 1677 ACTAAATATATTAAATGTAAGAAAGCCACCATTTCATGGGAAATATGCTAAGAAATTGG 1736
 OY ::::: |||
 OY 371 AlaGlnAlaMetAlaLysGlnGluHisThrGlnLeuLeuAlaGluIleLysGlnLeuHis 390
 Db 1736 ----- 1736
 OY 391 LysGluArgGluAspArgIleThrPheIleThrAsnIleValValAlaSerSerGlnGlu 410
 Db 1737 ATCGAAGACGAGATTTGTGTGAGAGCTGAGAAAC----- 1772
 OY ::::: |||
 OY 411 SerGlnGlnAspGlnArgValLysArgLysArgArgValThrThrAlaProGlyLysIle 430
 Db 1773 GCTGAAGAACTCAAAATGTGGAAACTAA----- 1802
 OY ::::: |||
 OY 431 GlnAsnSerLeuHisAlaSerGlyValSerAspPheAspMetLeuSerArgLeuProGly 450
 Db 1803 -----CTTCTTGATGAAGATCTAGATAAAACATTAGAGGAA 1838
 OY ::::: |||
 OY 451 Asn-----PheSerLysValAlaLysPheSerAspMetProSerPhePro 465
 Db 1839 AATAAGCCTTTCATTAGCCAGAGAGAGAAAGAAACTGTGTGACTTA----- 1886
 OY ::::: |||
 OY 466 GluIleAspAspSerValCysThrGluPheSerAspPheAspAlaLeuSerMetMet 485
 Db 1887 -----ATAGAAAGACTTGTGAAGAAAGAAAGAAAGATGAATGA 1919
 OY ::::: |||
 OY 486 AspSerAsnGlyIleAspAlaGluThrAsnLeuAlaSerLysValThrHisArgGluLys 505
 Db 1920 AAAAAAGAAATTAACCTTGAAATTAAATTCGAGAGAGATTAACACAGAGACTTACT 1979
 OY ::::: |||
 OY 506 ThrSerLeuHisGlnSerMetIleAspPheGlyGlnIleSerAspSerValGlnPheHis 525
 Db 1980 CAGTTATGGGCTCAACGGGAACTGACTTTAAGAG-----ACTCTGCTTCAAGAGACA 2033
 OY ::::: |||
 OY 526 AspSerSerLysGluAsnGlnLeuGlnThrLeuProLysAspSerGlyAspMetAlaGlu 545
 Db 2034 GAGATATTAGAGAAAT-----GCTGAA 2057
 OY ::::: |||
 OY 546 CysArgLysAlaSerPheGluLysGluIleThrSerLeuGlnGlnLeuGlnSerLys 565
 Db ----- 565

Db 2058 CCGCTTTGGCTATCTTCAG-----GATTTGGTTGGTAAATGTCAGACTGCA 2105
 OY ::::: |||
 OY 566 GluGlu-----GluLysLysGluLeuValGlnSer 575
 Db 2106 GAAAGAAAGCAGCGAAAGACATTGTGCCAAGAAAGTTGAACTGAAAGAGCTACTGCTGT 2165
 OY ::::: |||
 OY 576 PheGluLeuLysIleAlaGluLeuGluGluGlnLeuSerValLysAlaLysAsnLeuGlu 595
 Db 2166 TTGAAGCTAAATTTAAATCAATTTAAAGCTGAATTAGCT---AAACCCAAAGCA---GAA 2219
 OY ::::: |||
 OY 596 MetValThrAsnSerArgGluHisSerIleAsnAlaGluValGlnThrAspValGluLys 615
 Db 2220 TTAATC-----AAACCAAGCA 2237
 OY ::::: |||
 OY 616 GluValValArgLysGluMetSerValLeuGlyAspSerGlyTyrAsnAlaSerAsnSer 635
 Db 2238 GAGTTAAAGAGAGAA-----AATGAATCAGATTTCA 2270
 OY ::::: |||
 OY 636 AspLeuGlnAspSerSerValAspGlyLysArgLeuSerSerHisAspGluCysIle 655
 Db 2271 TTGATTCAGAGCTTGAGACATCTAATAAGAAATTAATTAACAGAAATCAAAATTTAA 2330
 OY ::::: |||
 OY 656 GluHisArgLysMetLeuGluGlnLysIleValAspLeuGluGluPheIleGluAsnLeu 675
 Db 2331 GAAATGATAAATATTAATGATGTCAGAAAGAAAGATACATCAAGAAATTT---CAGAACTCA 2387
 OY ::::: |||
 OY 676 AsnLysLysSerGluAsnAspLysGlnLysSerSerGluGlnAspPheMetIleSerIle 695
 Db 2388 AAGTCATATGGAACCATTTAAATGCAATGCAAGACAGCTGAT----- 2432
 OY ::::: |||
 OY 696 GlnLeuGluGluAlaIleMetAlaGluLysAlaAsnAlaLeuGluLeuAlaLeuMet 715
 Db 2433 ---ACATCTCTTAAATTAATTAATTAATTAATTTGTAATGAACAACTGTAAGTACCT 2489
 OY ::::: |||
 OY 716 ArgAspAsnPheAspAsnIleIleLeuGlnAsnGluThrLeu---LysArgGluIleAla 734
 Db 2490 AAGGACAGCAAACTAATAAATCTGTCAGAAAGAAAGAAAGATTAATGAATGAATCTGAC 2549
 OY ::::: |||
 OY 735 AspLeuGluArgSerLeuLysGluAsnGlnGluThrAsnGluPheGluIleLeuGluLys 754
 Db 2550 CAAGATGAACCCACAGCAAGAAAGAGGTATATCCATGTTAGTTCAGCTATACATGAAAGAC 2609
 OY ::::: |||
 OY 755 GluThrGlnLysGluHisGluAlaGlnLeuIleHisGluIleGlySerLeuLysLysLeu 774
 Db 2610 CAAAGAAAGAACTGAAGAAAGTCCGACCGACATTCAGAAATTTGAAGCAACAGATTTTA 2669
 OY ::::: |||
 OY 775 ValGluAsnAlaGluMetTyrAsnGlnAsnLeuGluLysPheGluThrLysThrLys 794
 Db 2670 CAAGAAAT-----AATGAAGAGCTGAGACATTTTACACTATTTAGAAAT 2717
 OY ::::: |||
 OY 795 LeuLeuLysGluGlnGluIleGlnLeuAlaGluLeuArgLysArgAlaAspAsnLeuGln 814
 Db 2718 GACTTTAAATTAAGAAAGAGAAAGAAAGCAATTAATTAACAGATTTGTTCAATTTTCAG 2777
 OY ::::: |||
 OY 815 LysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGluLysCysGlu 834
 Db 2778 CAG-----GAACCTTCTCTTCTGAAAAAGAAATTAACCTTTAAGTAA 2822
 OY ::::: |||
 OY 835 GluIlePheGlnLeuLysGlnSerLeuSerAspAlaGluAlaValThrArgAspAlaGln 854
 Db 2823 GAGGTCCACCAATTCAGTCGCAATTTATGATATTCGAATTCGT----- 2864
 OY ::::: |||
 OY 855 LysGluCysSerPheLeuArgSerGluAsnLeuGluLeuLysGluLysMetGluAspThr 874
 Db 2865 ---GAATTCAATGTCAGAAAGTAAAGTAAAGTAAAGCAAGGAGAAAGATATGAAATTTG 2921
 OY ::::: |||
 OY 875 SerAsnTyrArgAsnGlnLysGluLysAlaAlaSerLeuPheGluLysGlnLeuGluThr 894
 Db 2922 TCA-----AATGAGTTAGAAACGCTACACAGACATTTACAAATATGTTTCACAA 2972
 OY ::::: |||
 OY 895 GluLysSerAsnTyrLysMetGluAlaAspLeuGlnLysGluLeuGlnSerAlaPhe 914
 Db 2973 ATAAATTTATGTCACAGAAATTAAGCAAGTACGTAAGTCTGATTCAGTTTCAGATTT 3032
 OY ::::: |||

QY .915 AsnGluIleAsnTrpLeuAsnGlyLeuValGlyValProArgAspLeuSer 934
 Db 3033 TCACACATAGATTTCCTCAT----- 3053
 QY 935 ArgValGluLeuGluLysValSerGluPheSerLysGlnLeuGluLysAlaLeuGlu 954
 Db 3054 -----CTCAGGAGATCTGTCAAATGCTTCAG 3080
 QY 955 GluLysAsnAlaLeuGluAsnGluValThrCysLeuSerGluTrpLysPheLeuProAsn 974
 Db 3081 GAGGATAT-----TTGCCAAT 3098
 QY 975 ---GluValGluCysLeuLysAsnGln-----IleSerLysAlaSerGluLysLeuMet 991
 Db 3099 ACACAGTAGACCTTTAGGATATATATTTGGTAAAGCAAGTTAAACAAATATGCA 3158
 QY 992 LeuLeuLysGlnGluGluLysHisSerAlaSerIleIleSerLysGlnGluIleLeuMet 1011
 Db 3159 ATT-----CAAGAACCCCAATAGGAAATCTTCCACTAGATATGAAAGCTATTGG 3212
 QY 1012 GlnGluGlnSerGluGlnIleLeuGlnLeuThrAspGluValThrHisThrGlnSerLys 1031
 Db 3213 ---GAAAGATGTAAAGAGATTGTGAAGCCTTCCAAAAGCAATCAGATTGAGAA 3269
 QY 1032 Val---GlnGlnThrGlnGluGlnIleTrpLeuGluMetLysMetHisAspSer----- 1048
 Db 3270 CTGCAACACAAATTTGAAAATTCGACGACAGATGAAAAGCTATTAAGATGAAACAAAT 3329
 QY 1049 ---LeupheGluLysTrpIleArgAsnLysSerGluValGluAspLeuLeuArgLysLeuMet 1067
 Db 3330 AGACATTAAGGAGAAAGGACATTAATAACCA-----GATGACCTCTTAATAAATAA 3380
 QY 1068 GluAsnLeuLysGlyThrMetGluSerValGluValLysIleAlaAspThrLysHisGlu 1087
 Db 3381 GAACCTCT-----ATACACACCTGAAAGAAAGAA 3410
 QY 1088 LeuGluGlu---ThrIleArgAspLysGluGlnLeuLeuHisGluLysTrpPhePhe 1106
 Db 3411 TTGCAGAAAATAATGTACCTTGATGTTCAATACAGCAT----- 3452
 QY 1107 GlnAlaMetGlnThrIlePheProIleThrProLeuSerAspSerLeuProProSerLys 1126
 Db 3452 ----- 3452
 QY 1127 LeuValGluGluLysAsnSerGlnAspProIleGluIleAsnAspTrpHisAsnLeuIleAla 1146
 Db 3453 GTAGTTGAAGGAAAG-----AGA 3470
 QY 1147 LeuAlaThrGluArgAsnAsnIleMetValCysLeuGluGluThrGluArgAsnSerLeuLys 1166
 Db 3471 GCGCTTCAGAACTTACACAAAGGTGTACTTGC-----TATTAAG 3509
 QY 1167 GluGlnValIleAspLeuAsnThrGlnLeuGlnSerLeuGlnAlaGlnSerIleGluLys 1186
 Db 3510 GCAAAAATAAAGGAACTTGAAGAA-----ATTTTAGACTCACAAGAAAGTGAAGCT 3560
 QY 1187 SerAspLeuGlnLysProLysGlnAspLeuGluGluGluGlu-----ValLysLeu 1203
 Db 3561 AGTATTCACGCCAAGTTAGAACAGACATTTTGGAAAAGAAATCTATCATCTTAAAGCTA 3620
 QY 1204 LeuLeuGluMetGluLeuLeuLysGlyHisLeuThrAspSer-----Gln 1218
 Db 3621 GAAAGAAATTTGAAGGAATTTCAAGAACATCTTCAGATCTGTGCAAAAACACCAAAAGAT 3680
 QY 1219 LeuSerIleGluLysLeuGlnLeuGlnLysAsnLeuGluValThrGluLysLeuGlnThrLeu 1238
 Db 3681 TTAATATGAAAGCAATTCACCTGAAAGAA-----GAAATCACACAGTTTAACAAATTAATTG 3737
 QY 1239 GlnGluGluMetLysAsnIleThrIleGluArgAsnGluLeuGlnThrAsnPheGluAsp 1258
 Db 3738 CAAGATATGAACATTTACTTCAATTAAAGACAGAAAGAAAGAACCAAC----- 3788

QY 1259 LeuLysAlaGlnHisAspSerLeuLysGlnAspLeuSerGluAsnIleGluGlnSerIle 1278
 Db 3789 ---AGCCAGAAAGACAGAAAATTTGAAAGAGAACTCTGCAAGCTCTGTCACCCAG 3845
 QY 1279 GluThrGlnAspGluLeuArgAlaAlaGlnGluGluLeuArgGlnGlnLeuVal 1298
 Db 3846 AATCTGAAAGCAGATCTTCAGAGAGAAAGAAATTAATGCTGACCTGAAAGAAACTG 3905
 QY 1299 AspSerPheArgGlnGlnLeuLeuAspCysSerValGlyIleSerSerProAsnHisAsp 1318
 Db 3906 ACTGATGCCAAAACACAGAT----- 3926
 QY 1319 AlaValAlaAsnGlnGluLysValSerLeuGlyGluValAsnSerLeuGlnSerLeuMet 1338
 Db 3926 ----- 3926
 QY 1339 LeuArgGlyGluArgAspGluLeuGlnThrSerCysLysAlaLeuValSerGluLeuGlu 1358
 Db 3927 -----AAGCAATACAGAAAGAGATATCTGTAATGCTGATGAGATTA 3971
 QY 1359 LeuLeuArgAlaHisValLysSerValGluGlyGluAsnLeuGluIleThrLysLysLeu 1378
 Db 3972 TTACTGAGGATT-----AAAAT 3989
 QY 1379 AsnGlyLeuGluLysGluIleLeuGlyLysSerGluGluSerGluValLeuLysSerMet 1398
 Db 3990 AATGAATCTGAGAAAAGAAAAGAAACCAAGTCTTCAGGAATTAAGATTAAGACAGCAAC 4049
 QY 1399 LeuGluAsnLeuLysGlu-----AspAsnAsnLysLeuLysGluGlnAlaGluGluTrp 1416
 Db 4050 ATTCAAGCAATCTCAAGAGCAGTTAAATATCAAGAAAGTGAAGAGCTATACACAGTAT 4109
 QY 1417 SerSerLysGluAsnGlnPheSerLeuGlnValPheSerClySerGlnLysLeuVal 1436
 Db 4110 GACAGACATCGAAAAGATTAATGTTAAAG-----AAAATTAAT 4151
 QY 1437 AspGluIleGluValLeuLysAlaGlnLeuLysAlaAlaGluGluArgLeuGluLysLeu 1456
 Db 4152 GAAGACATG-----CGAATGACACTAGAA 4175
 QY 1457 AspArgAspTrpPheGluLeuValGlnThrAlaAsnThrAsnLeuValGluLysLeu 1476
 Db 4176 GACAG-----GAACAACTCAGGTAGAACAGATCAACAGCTTCGACGCTAAATTA 4226
 QY 1477 GluThrProLeuGlnAlaAspHisGluAspSerIleAspArgSerLeuGluMet 1496
 Db 4227 GAG----- 4229
 QY 1497 GluIleLysValLeuGlyGluLysLeuGluArgAsnGlnTrpLeuLeuGluLysLeuGln 1516
 Db 4230 GAAGTTGAAAGCTCTGCCACAGAAATTTGAAAATGGAAGAAATGCAATGATTTGGAA 4289
 QY 1517 GluGluLysLeuGluLeuSerAsnLysLeuGluIleLeuGlnLysGluMetGluThrSer 1536
 Db 4290 ACCAAAACAAATCAAAAGTCAAAATTA-----GAACATGACAGAACACACAGAT 4337
 QY 1537 ValLeu-----LeuLysAspAspLeuGlnGlnLysLeuGlnSerLeu----- 1550
 Db 4338 GTGCTTGAAGACCTCACTAATCTTCAAGATGAGTTACAGAGATCTGAACAGAAATATATAT 4397
 QY 1551 -----LeuSerGluAsnIleIleLeu-----Lys 1558
 Db 4398 GGTATAGAAAGAAATGGTTAGAAAGAAAATGATGCTTATACATCAACGCAAGAAAGACA 4457
 QY 1559 GluAsnIle---AspThrThrLeuLysHisHisSerAspThrGlnAlaGlnLeuLys 1577
 Db 4458 GAGATATATCGAAATTAAGAGATGTAATAAATATGCTGAGACAGGACGCTTTTATTAAG 4517
 QY 1578 ThrGlnGlnGluLeuGlnLeuAlaLysAsnLeuAlaIleAlaLysAspAsnPro 1597
 Db 4518 CAACGATATGAATATGAA----- 4535
 QY 1598 IleThrGlnGluLysGluThrSerAlaAspCysValHisProLeuGluGluLysIleLeu 1617

Db 2413 GATTGGATCAAAATACCAAAATTTCTAACCACTA---GATTTTCTTAAACGT 2469
Qy 1421 AsnGlnPhe-----SerLeuGluGlu 1427
Db 2470 TATGAATATGCTCGAAGATATGTTGAAGATATGCTGAGAAATACATCTACCTCATGAG 2529
Qy 1428 -----ValPheSerGlySerGlnLysLeuValAspGluIleGluValLeuLys 1443
Db 2530 AGAAATACGAAATCTACCTGAGCAATCAAAAGCAAGAACAGATTAATCAATACATGACT 2589
Qy 1444 AlaGlnLeuLysAlaAlaGluGluArgLeuGluIleLysAspArgAspTyrPheGluLeu 1463
Db 2590 CAAGATTGTGAGAGAGCAATAGAGAGCTGCTGCGAGA--- 2631
Qy 1464 ValGlnThrAlaAsnThrAsnLeuValGluLysLeu-----GluThrProLeu 1480
Db 2632 GTAAAGCAGAAATTTGAGAGAGAAAGAAATCTTAATTTGTCGAGATTCGCTT 2691
Qy 1481 GlnAlaAspHisGluGluAspSerIleAspArgArgSerGluGluMetGluIleLysVal 1500
Db 2692 TCTCAGCAAAAGAGCTCTTTGTTTACCTGACAAAGGGGCAAACTTA--- 2739
Qy 1501 LeuGlyGluLysLeuGluLysArgAsnGlnLysLeuGluArgLeuGln---GluGluLys 1519
Db 2740 CTGCTAATCAATCTGCAAAACATTCAGGAAATCTGAGCGATCTGAAACAGAAACAA 2799
Qy 1520 LeuGluLeuSerAsnLysLeuGluIleLeuGlnLysGluMetGluThrSerValLeuLeu 1539
Db 2800 CAAAGCTTACAGTACGACAGTACGAAACACTGCAACATGAGATC--- 2841
Qy 1540 LysAspAspLeuGlnGlnLysLeuGlnLysLeuSerLeuSerGluAsnIleLeuLysGlu 1559
Db 2842 ---TCTCATCTTAAGAGAGAGTTGCAAAATGAGTGCAACAAAGCATCACTTCTAGA 2898
Qy 1560 AsnIleAspThrThrLeuLysHisSerAspThrGlnAlaGlnLeuGlnLysThrGln 1579
Db 2899 AATCTAGAT-----GTTCAACTTTTACATCAAG 2928
Qy 1580 GlnGluLeuGlnLeuAlaLysAsnLeuAlaIleAlaSerAspAsnProIleThr 1599
Db 2929 AGCAACTGAGTACAGAGCAAACTTT--- 2955
Qy 1600 GlnGluLysGluThrSerAlaAspCysValHisProLeuGluGluLysIleLeuLeu 1619
Db 2956 -----CAT---CTTAAC 2964
Qy 1620 ThrGlnGluLeuHisGlnLysThrAsnGlnGlnLysLeuHisGluLysAsnGlu 1639
Db 2965 ACAAAAGAACTATTAA---AATGCTCAAAAGAAATGCTCCATGAAACAGCAC 3018
Qy 1640 LeuGluGlnAlaGlnValGluLeuLysCysGlu---ValGlnHisLeuMetLysSerMet 1658
Db 3019 CTCAGTAAATGAGAGCCAAAGTGTCTGCTTCCACAGAGAACTGTAAGAGTCAAG 3078
Qy 1659 IleGluSerLysSerSerLeuGlnSerLeuGlnHisGluLysHisAspThrGlnGln 1678
Db 3079 CCTAGCAACAAAGAGATGTGATGATCTTGTAGTCAGCTAAGACAGACAGAGAGCAG 3138
Qy 1679 LeuLeuAlaLeuLysGlnGlnMetGlnValIleThrGlnLysLysGluLeuGln--- 1697
Db 3139 GTGAAATGACTTAAAGGAGAACTCAAAACAAAGTACGAAATGTAACAAATTCAGAA 3198
Qy 1698 -----GlnThrHisGlnHisLeuThrAlaGlu---ValAspHisLeuLys 1711
Db 3199 ATGCTTACTAGTTAGAAAGAAATCCCTGAAACAGAAAGAGTGCACAGAAAGTGCCT 3258
Qy 1712 GlnAsnIleGluLysGlnLysAsnPheLysAsnGlnAlaGlnGlnLysThrThrLysGlu 1731
Db 3259 AAGAAATTTGAAAGTCTGTTA---AAGAGTCACTTAATTCAGACAGCTTGAA 3312
Qy 1732 GlnLysLeuLeuAsnGlnLysGlnLysGluLeuGlnGlnSerGlnHisArgLeuGlnCysGlu 1751
Db 3313 AAGAAATTTGAGTGAAGTGAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTCAAG 3354

Qy 1752 IleGluGluLeuMetLysSerLeuLysAspLysGlnSerAlaLeuGluIleThrLysGlu 1771
Db 3355 -----CATGATTAAGAGAGCCATGAGAGCATG--- 3384
Qy 1772 SerGlnLysValIleAsnLeuAsnGlnGlnMetGluMetValMetLeuGluMetGlu 1791
Db 3385 ---GAACAACAGTTATCTGAAATGAAAGAAACACTTCTAGTGTTCAGAAATGATCAAC 3441
Qy 1792 GlnLeuLysAsnSerGlnArgThrValIleAlaGluArgAspGlnLeuGlnAspLeu 1811
Db 3442 GAAGCTTCTCAGAGACCAAGCAAGCTTAAAGATGATGACACCAAGCCAGAGCTACGT 3501
Qy 1812 ArgGlnSerValGlnMetSerIleGluThrGlnAspAspLeuArgLysAla----- 1828
Db 3502 CAGGAACAGCTAAATTAATGCTGCAAGCTCAAGATTAAGTACAGAGAAATGATGCTG 3561
Qy 1829 -----GlnGluAlaLeuGlnGlnLysAspLysValGlnLysLeuThrSer 1844
Db 3562 CATGCTGCTGATGTTGAAAGCTTACAAAGCTGCAAGAGAGAGAGTTCAAATGAGCATCA 3621
Qy 1845 GlnIleSerValLeuGlnGlnLysIleSerLeuLeuGlnAsnGlnMetLeuThrAsnVal 1864
Db 3622 GTCCGTCAGCATTTGGAAGAAACACACAGAAAGCAAGATCAAGTGTGGAGTAA 3681
Qy 1865 AlaThr-----ValLysGluThrLeuSerGlu 1873
Db 3682 GCATCTTGGGAGGAAAGAGAGAAATGTTAAAGATGAAGTTCCAAATGTGATGCTGC 3741
Qy 1874 ArgAspAspLeuAsnGlnSerLysGlnHisLeuPheSerGluIleGluThrLysSer 1892
Db 3742 TGTAAGATCTGGAGAAACAAACAGATTTCTTCATGATCAGATCCAAATTAATGATGAC 3801
Qy 1893 -----LeuSerLeu----- 1895
Db 3802 AAGTGTGCTGCTGTGGAAGAGAGTGTACAAAGTCCATGATGATGCTCACTGTA 3861
Qy 1896 -----LysGluLysGlu 1899
Db 3862 GAAGAAATCTCAAGAAACAAATTTGGAATTTCTCAGATTTATACGACAGAAAGAA 3921
Qy 1900 -----PheAlaLeuGlnAlaGlu--- 1906
Db 3922 ATGCTGAAACTAGATTTGAGTGTGCTGAGTGTGAGAGTGTGCTGATTCGAAAGGCT 3981
Qy 1907 -----LysAspLysAlaAspAlaAlaArgLysThr 1916
Db 3982 GAATCTTTAGAAAGAGAGCTCCAGAGCTCGAAAGATGCTTAATGCTGAAGAGGAGAA 4041
Qy 1917 IleAspIleThrGluLysIleSerAsnIleGlnGlnLeuGlnGlnAlaThrAsn 1936
Db 4042 GTCCAGTAACTGCAAAACAAATGGCTCAGCATGACAGAACTGATGAAGAAACTGAACA 4101
Qy 1937 LeuLysGluThrLeuThrGluArgGluSerLeuIleGlnCysLysGlnGlnAlaLeu 1956
Db 4102 ATGATGTAGTTATGAGAACCAATAAATGCTTAAGAGAGAGAGAGAGAGCTGAACAG 4161
Qy 1957 AsnThrGlnHisLeuArgGluThrLeuLysSerLysAspLeuAlaLeuGluLysMetGlu 1976
Db 4162 GATCTACAGAAATCTCAAGCAAGAGTGAAGAACTGAGATTAATTTACCTTACAA 4221
Qy 1977 GlnGluArgAspGlnAlaAlaAsnLysValIleAlaLeuThrGluLysMetSerLeu 1996
Db 4222 GAAGCAATGCTGAGCTGAGTGAAGAAAGCGTATGTCAGAGCAAGAAAGCTCTTA 4281
Qy 1997 GlnGluGlnIle-----AsnGlnAsnValIleThrLeuLysGluGlu 2010
Db 4282 GAAGAGAGTGTCAAAAGTGTGAAGAACAGCTAACAGCATCTAGTAAGI----- 4329
Qy 2011 GlnGlyGlnLysGluThrPheTyrLeuGlnArgProSerLysGlnGlnSerSerGln 2030
Db 4330 -----CAACAGAAAGATCCAGAT 4347

QY 2031 MetGIuIleuArgIleuSerIleuLysThrLysAspIleuGluIleuGluIleuValIleuLys 2050
 4348 ACAGAAAGATATGCGAAGCCCTTCTGAAAAGGAAGTTAT-----ACTAAG 4395
 QY 2051 GluIleSerGluIleThrAsnGluIleLysAsnIleuThrAlaLysIleSerSerIleuGlu 2070
 4396 CGTATTCACAAATTGACAGAAATTTGGTAGACTTAAGCTGAATTCGAATGCAATCAAT 4455
 QY 2071 GluGluIleuGluIleuAsnIleuSerIleuLeuAsnIleuValIleuSerGluIleuAsnIleu 2090
 4456 GCATCTTCTACTAACACAGAACTTAATTCAGAGCTGTGAAGGAAGTCTAAATTAAGTA 4515
 QY 2091 ArgHisSerLysGluIleuValSerGluIleuGluIleuSerIleuThrIleuLysSer 2110
 4516 AGAAGCTGAAGAAAGCAACCTCCAGAGAGACTTGATCCCAATTAATGTATCCAGAA 4575
 QY 2111 ArgAspHisAlaPheAlaGlnSerLysArgGluIleuAspGluIleuValIleuLysIleAla 2130
 4576 AAGGTCAAAACTATTTACTCAAGTTAAGAAATTTGACGTAGTACAAAGACTCAATATGAA 4635
 QY 2131 SerIleu-----AlaGluIleuIleuLysIleuThrLysGluMetAspGluPheArgAspSer 2149
 4636 GAACCTTAAGACACACAGATTAAGTTATGGAGACATCGGCTCAGTCTCTGGAGACCAT 4695
 QY 2150 LysGlu-----SerIleuGluIleuGlnSerSerHisIleuSerGluIleuLysThr 2166
 4696 CAGGACACAGATGTTTCAGTCCAGGAATG---CAGGAAGCTCAAGAAAGCGTCAACCAA 4752
 QY 2167 TyrLysThrGluIleuGlnMetLeuLysGluIleuGluIleuAspIleuAsnIleuLysAla 2186
 4753 GGTGAACCAAAATCAAAATCACTTGAAGTCAAGTACAGATCTGCAAGAACATTAATCT 4812
 QY 2187 GluLysValIleuGluIleuValIleuAspGluIleuGlnHisIleuSerSerIleuGluIleu 2206
 4813 GAAAGAAAGACACAGAACCAAAATCTCCAGAAACAGACTGTGCAACTTCACTGTAACCTT 4872
 QY 2207 AspGlnIleGlnMetGluIleuArgAsnGluIleuLysArgAsnIleuLysGluIleuLys 2226
 4873 TCACGACCTTCGTACAGATGTTCAAGATTAAGACACACAGAGAGAGACAGCTC-----CGA 4926
 QY 2227 MetAspIleuMetGluIleuGlu----- 2233
 4927 CACACATATCACTGAAGAAAGAAAGAAACCAGAAAGCTATTGTACAGCAAAAGTCAAA 4986
 QY 2234 IleSerValIleuArgIleuMetGlnAsnGluProGluIleuGluIleuAspValIleuGlu 2253
 4987 ATTGCACACTTACCTGCTGTAAAAGATCACTTAAGTAAAGAAATGAGAGCTTAACAA 5046
 QY 2254 ArgMetAspIleuGluIleuSerArgAsnGluIleuGlnIleuMetGluIleuLysIleSer 2273
 5047 AGGAATGAGACCTTACAT-----CAGCAAGAAAGATGAATTTGGATGTTCACATTACT 5097
 QY 2274 AlaValIleuSerGluIleuHisThrIleuLeuSerSerIleuSerSerGluIleuGluIleu 2293
 5098 GCGCTAAAGTCCCAATATGAAGGTCAAGTTCAGCTTGAAGAAAGCAACACTCAGGAGCAT 5157
 QY 2294 ThrGluAlaHisIleuLysIleuSerMetLeuAsnIleuLysGluIleuSerSerThrIleu 2313
 5158 CAAGAGACAC-----CTTGACACAGAGATGAGCTCAGCAAGCTTCTTAAT 5205
 QY 2313 ArgSerPheGlySerIleuGlnThrGluIleuHisValIleuLysAsnThrGluIleuGlnThrIle 2333
 5206 AAGGTCCCGAAGACAGACAGATGATGCAATTCAAA-----CAACTCCAGCTTC 5256
 QY 2333 IleuAsnIleuLysValIleuValIleuArgThrAlaAlaValIleuGluAspHisSerIleuI 2353
 5257 G-----GTCAAGAGAC----- 5266
 QY 2353 LysAspTyrGluLysAspLeuAlaIleuGlnLysArgHisAspIleuArgIleuGlu 2373
 5267 -----GAAATTCACACATATCAAGCCACCAACAGCAATATCAAGCCA 5310
 QY 2373 n---IleuGlnCysIleuGluIleuHis-----GlyArgLysTrp 2384

Db 5311 ACTGCTGTGTGCTACTCCAAAT-AAAATGACAGCTGCAGCATATGGCTGGAATAATATGTC 5369
 QY 2384 pSerAspSerAlaSerGluIleuLysPheCysGluIleGluPheLeuAsnGluIleu 2404
 5370 AACACCCAGGCGTATTCGCCCAATGTTTACACCTGCACAACTGTTCAAAATCCCACTAC 5429
 QY 2404 uPheLysLysAlaAsnIleuIleuGlnSerValGlnAspAspPheSerIleuValGlnValPhe 2424
 5430 TACCCCAACAGCTACAGTGAATGCCACTACACAAAGTGAA----- 5469
 QY 2424 IleuAsnGlnValGlySerThrIleuGlnGluIleuGlnHisLysLysGlyPheMetGlu 2444
 5470 -----TCACAGAAAGCTATGCAGTACAGCAAGAGCGCTGTGGA 5504
 QY 2444 nTyrPheGluIleuPheGlyLysPheHisValAspAlaLysLysLeuSerGluIleuMetGlu 2464
 5505 ACATGTTCAGTGTTCATTCAGCAACAGTGCCTGTCCTTCACTATGTCCTCA 5564
 QY 2464 nGlnGluAsnArgIleuAlaSerThrIleuGlnIleuLeuThrLysArgLeuLysAlaVal 2484
 5565 GCCTTCATCTCCTACACTTATTTAACTGTTCACACAAACA---CAGGCTACAGCTTT 5621
 QY 2484 ValGlnSerLysIleGlnArgGluIleuThrValTyrLeuAsnGlnPheGluIleuLysLe 2504
 5622 TGTGCAACCCACATCAACAGATCAT-----CTTCAGATTCAGCCTGCGCA 5666
 QY 2504 uGlnGluLysLysGluIleuAsnLysGluIleuMetArgArgMetGlnHisIleuProse 2524
 5667 TCAAGATTAATCTTCAAAACATAGTAGAGGTTCACAGATTCACACCTTACGCGCTTC 5726
 QY 2524 PAlaSerValIleuGluIleuGluIleuAsnAlaArgLeuGluIleuLysThrValGlnAs 2544
 5727 TACTTCC-----ACAGCAGATTTTGGCAGACATTCGGCTACCCCGAG 5768
 QY 2544 pGluSerLysLysLeuGlnSerArgIleuLysMetLeuGluAsnGluIleuAsnLeuValLys 2564
 5769 TCTCTTCTTCCCAAGCGTACAGCT-----GAAGAGAGA----- 5802
 QY 2564 sAspAspAlaMetHisLysGluIleuLysValAlaIleuGlnAspLysLeuLeuSerArg 2584
 5803 -----GAGATGACACCATAGAAC 5822
 QY 2584 gAsnAlaGluAlaGluIleuAsnAlaMetGlnValLysLeuThrLysLysGlnAspLeu 2604
 5823 ATCAGACCAAGTCTGTATGATATACAGTGAATGCCCTTCCAAAGAG----- 5871
 QY 2604 uGlnAlaAlaMetLysGluIleuGluIleuGlnLysMetValAlaLysGluIleuValPhe 2624
 5872 -----TTGAAAAGTGCACACCC 5888
 QY 2624 cTyrLysGluIleuIleuAspAsnLeuLysThrLysValIleuValLysIleuMetGluLysIle 2644
 5889 TGTAGCAACTGAGGAAGAAATTTATGCCAGAAAGAAAGTATGATGAGAGGTAGAGACTCA 5948
 QY 2644 eLysTyrSerLysAlaThrAspGlnGluIleuAla-----TyrLe 2657
 5949 GGTATACACACAGATTCCTCAAGTTCATTTGGAGAAAGAGTATACCCAGGAGATTTAT 6006
 QY 2657 uLysSerCysLeuGluIleuAspLysGluIleuLysLeuArgArgLeuLysGluIleuArgArg 2677
 6007 -----ACACCTATGGAAGACAGTGAAGAAACCTCTCACTGCTCAAAATATGATCTTGGGCC 6062
 QY 2677 gAlaGlnAlaAspAsnAspThrThr----- 2685
 6063 ACTTCAATCAGATCAGACGACGACCACTTCATCCAGAGATGTCACGAAGCAAGAGATGA 6122
 QY 2686 ---ValCysValProLysAspTyrGlnLysAla----- 2695
 6123 TGTCAATTGTTATTTGACAGTATGATGTAAGAAAGAGATGAGAAAGATGATGATGATCA 6182
 QY 2695 ----- 2695

```

Db 6183 AGATGACACAGGATGGAGATGAGGTGAAGATAGTAAAGAACTGGTAGTGGCGA 6242
Qy 2696 -----SerThrpheProValTh 2701
Db 6243 TGGCAATGATGATTGACATGATGCTGAGGCTGATGAGGAGCTGATCCAGTAC 6302
Qy 2701 T-----CysGlyGlySerGlyIleValGlnSerThrAlaMetLeuVa 2716
Db 6303 AGAAACAGAAAGAAAGTATGGGTGAGGTGAAGT----- 6336
Qy 2716 IleuGlnSerGlyAlaIleuGlnArgGluLeuSerHisTyrLysLysTyrH 2736
Db 6337 -----AATCACAAGAGCTGATGATCTCAAAACAGTGTGAAGAAATACAGTCTGC 6389
Qy 2736 SHISLeuSerArgThrMetSerSerSerGluAspArgLysThrLysAlaLysSerAs 2756
Db 6390 AGAATCTTTCTTTCTCAGAGAGTTCTAGAGAAACAAGCCATCATACACA---TCTGA 6446
Qy 2756 pAlaHisSerSerHisThrGlySerSerHisArgGlySerProHisLysThrGluThrTy 2776
Db 6447 AAGACAGGCCCTCGACACCTCAGACCGAGACGCCACACACAT----- 6492
Qy 2776 rArgHisGlyProValThrProGluArgSer---GluMetProSerLeuHisLeuGlySe 2795
Db 6493 -----CCACTTCCCCCAAGACTGACCAATTCATGCCCACTCAGAGAGTGGACC 6542
Qy 2795 rProLysLysSerGluSerSerThrLysArgValAlaSerProAsnArgSerGluIleTy 2815
Db 6543 ACCA-----GTTCAAGAAATTCAGATGACGCCAGCCAGCACTGTAGG 6584
Qy 2815 rSerGlnLeuValMetSerProGlyLysThrGlyMetHisLysHisIleLeuSer----- 2833
Db 6585 AGCTGGCCCTTCAGTGTACTCCAGAAATGAGTGGCAGTCAACAGCATTTTGTGATGATGA 6644
Qy 2834 -----ProSerLysValGlyLeuHisLysLysArgAlaLeuSerProAsnArg 2849
Db 6645 AGACAGACAGTTCACAAAGTACTCCAACTCT-----GTGGTCCACATCG 6689
Qy 2849 gSerGluMetProThrGlnHisValIleSerProGlyLysThrGlyLeuHisLys----- 2867
Db 6690 TACTATGATGATTTGCTGAAGCAATTCATTCGCCGAGGTGTGCTGTGCTCCAGATTCGG 6749
Qy 2868 -----AsnLeuThrGluSerThrLeuPheAspAsnLeuSerSerProCy 2882
Db 6750 GTTGGCCACCTGAAGATATGCCACAAACAAGT-----TCTATCTACTGATCT 6800
Qy 2882 sLysGlnGlnLysValGlnGlnAsnLeuAsnSerProLysGlyLysLeuPhe----- 2899
Db 6801 TGGCAGCTTCTCTCTCAAGAGGTTTAGAATGATATAAACACCCCTGTCTCTAGTCA 6860
Qy 2900 -----AspValLysSerLysSerMetProTyrCysProSerGln----- 2912
Db 6861 TGAAGAAAGATGAGTGGCCGAAGTGTCCACACTACTCATTACAAAGTAGCAGCCCCAGT 6920
Qy 2913 -----PhePheAspAsnSerLysLeuGlyAspPheSerGlu 2924
Db 6921 GACTGTATTACTGAGACACCACCTCTGATGCTTCGAA 6960

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 2171870CB1
US-10-044-090-359

Alignment Scores:
Pred. No.: 4.98e-48
Score: 891.50
Percent Similarity: 37.15%
Best Local Similarity: 20.35%
Query Match: 6.04%
DB: 12
Gaps: 97

US-09-150-867-1 (1-2954) x US-10-044-090-359 (1-7792)
Qy 600 SerArgGluHisSerIleAsnAlaGluValGlnThrAspValGluGlyAlaValArg 619
Db 111 TCCCGCCCGGAGTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 170
Qy 620 LysGluMetSerValLeuGlyAspSerGly-----TyrAsnAlaSerAsnSer 635
Db 171 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 230
Qy 636 AspLeuGlnAspSerSerValAspGlyLysArg---LeuSerSerSerHisAspGluCys 654
Db 231 TCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 290
Qy 655 IleGluHisArgLysMet-----LeuGluGlnLys 664
Db 291 CTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 350
Qy 665 IleValAspLeuGluGluPheIleGluAsnLeuAsnLysLysSerGluAsnAspLysGln 684
Db 351 CTGCTGATTCACGATTCGAGATGATGCTGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 410
Qy 685 LysSerSerGluGlnAspPheMetGluSerIleGlnLeuGluAlaIleMetAlaGlu 704
Db 411 GAGAGCGGCAACAGTATTT----- 431
Qy 705 LysAlaAsnAlaLeuGluGluLeuAlaLeuMetArgAspAsnPheAspAsnIleLeu 724
Db 432 -----GAATAGAAAGAGGTGTCTCCACAGTACAGAGAGACTTGAAATGAACCCGA 485
Qy 725 GluAsnGluThrLeuLysArgGluIleAlaAspSerLeuArgSerLysGluAsnGln 744
Db 486 GAGTGTCAAGCTTGGCGGCTTGAGCTAGAGAACTCAACATCACTGAAGGCACTACT 545
Qy 745 GluThrAsn---GluPheGluIleLeuGlu----- 753
Db 546 GAGAAACAAAGAACTTGAATTCGTCAGAGATCGCAATATTCATTCAGACCAATTT 605
Qy 754 LysGluThrGlnLysGluHisGluAlaGlnIleuIleHisGluIleGlySerLeuLysLys 773
Db 606 ACAAGAAACAAAGAGATTAAGAACTGAG-----AAAAGAGC 644
Qy 774 LeuValGluAsnAlaGluMetTyrAsnGlnAsnLeuGlu-----GluAspLeuGlu 790
Db 645 TTAATTAGAACCAATGAGACATATCTCAAGACTTGAAATTAATTACAGAGGATGTATAA 704
Qy 791 ThrLysThrLysLeuLysGluGlnGluIleGlnLeuAlaGluLeuArgLysArgAla 810
Db 705 CGTCTGATGACAAACCTTAAGAAACCAATACAAACAAGGCGTGAAGTTAAATTAATGG 764
Qy 811 AspAsnLeuGlnLysLysValArgAsnPheAspLeuSerValSerMetGlyAspSerGlu 830
Db 765 GATGAACCTCAAGCTCT-----GATGTTCTGTTAAGTATCGAGAA 806
Qy 831 LysLeuGluGluGlu-----IlePheGln 838
Db 807 AAACGCTTGAGACGAAAGAAAGAAATTCATACATAGTCAAGATACATGCTGAATACAGAG 866
Qy 839 LeuLysGlnSerLeuSerAspAlaGluAlaValThrArgAspAlaGlnLysGluCysSer 858

```

```

Db 867 TTGAAGAACCAAACTGATGAACTTCTGCTTGGAGAGAGAAAAGGAAATGAGAT--- 923
Oy 869 PheLeuArgSerGluAsnLeuGluLeuGluGluMetGluAspThrSerAsnTrpIyr 878
Db 924 -----CTAGAGCTTAATGTAATCTTGA--- 947
Oy 879 AangLysGluValAlaAlaSerLeuPheGluGluGluGluThrGluLysSerAsn 898
Db 948 AATAAAGAGAGAGAGGTTCT-----AGACTGGAGACAAATGAAATGGAGCC 992
Oy 899 TyrLysMetGluAlaAspLeuGluGluGluGluGluGluGluGluGluGluGluGlu 918
Db 993 TTAATAACATCAATGACATCTTCAAAAGCATGTGAG----- 1031
Oy 919 TyrLeuAsnGlyLeuLeuAlaGlyLysValProArgAspLeuSerArgValGluLeu 938
Db 1032 -----GATCTGTGTGACCAATTA----- 1049
Oy 939 GluLysValSerGluPheSerLysGluLeuGluLysAlaLeuGluGluLys----- 956
Db 1050 -----AAAGAGCCAAAGACACAGCCAGTATGGAGAGAAATTCAC 1094
Oy 957 AsnAlaLeuGluAsnGluValThrCysLeuSerGluTyrLysPheLeuProAsnGluVal 976
Db 1095 AATGAATTAATGCCCACATTAATCTTCTAATTTGTACAGAGTCCCGCTGATGACTCA 1154
Oy 977 GluCysLeuLysAsnGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 995
Db 1155 GAAGCAAGAGCAATGAACATAACCCGCGAGTGAAGAACTCACAAATTTTGAAGAA 1214
Oy 996 GluGluGlu-----HisSerAlaSerIleLeuSerLysGluGlu 1008
Db 1215 GCTGTGAAGCCAAAGCAATACAGATCATCTTCTAGAGTGGAGCAATCCAAAGAT 1274
Oy 1009 IleIleMetGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1028
Db 1275 CAATGGAGAAAGAAATGCTTGAAGAAATAGGAGATTTGAGAGAAATTAAGAAATGCA 1334
Oy 1029 GlnSerLysValGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1040
Db 1335 AATGACCTTCTTGTGCCCAAAAGCAATGAAGCCATATTTGTGAAAGAGCTTGGCC 1394
Oy 1041 GluMetLys-----LysMetHis 1046
Db 1395 GCCATGCTCTCTACTGACAGCACTGTAGTAACTGAACTGGAGTAACTA--- 1451
Oy 1047 AspAspLeuPheGluLysTyrIleArgAsnLysSerGluAlaGluAspLeuArgGlu 1066
Db 1452 ACTGAGCTCTAATGCTTATGTGAAACTCAGATCACTG-----CTTTGGAGAA 1505
Oy 1067 MetGluAsn-----LeuLysGlyThrMetGluSerValGluValLys 1080
Db 1506 CTAGAGAAACAAAGAAATTAATAGTACATAGATGAATGAATGAAGAAATGGAAGCCAA 1565
Oy 1081 IleAlaAspThrLysHisGluLeuGluGluGluGluGluGluGluGluGluGluGlu 1100
Db 1566 GCACCAATTTTGAAGCCAGCGTGAAGATATGAACGCGCAAGAACTGTAGCAACT 1625
Oy 1101 GluLysLysTyrPhePheGluAlaMetGluGluGluGluGluGluGluGluGluGlu 1120
Db 1626 TTATCTGTTAAGCTGACAAAGCTATGAAGAGATTCAGCGATGACAGAGACACTGAT 1685
Oy 1121 SerLeuProPheSerLysLeuValGluGluGluGluGluGluGluGluGluGluGlu 1140
Db 1686 -----AAAGCCAAAGCAATCATCTGTA----- 1709
Oy 1141 TyrHisAsnLeuIleAlaLeuAlaThrGluArgAsnAsnIleMetValCysLeuGluThr 1160
Db 1710 -----CTTGAGACA 1718
Oy 1161 GluArgAsnSerLeuLysGluGluGluGluGluGluGluGluGluGluGluGluGlu 1180

```

```

Db 1719 GATTAATGAAGAAATGGAATATCAATGAATGAATCTTTCACACAGATTAAGTCT--- 1775
Oy 1181 AlaGlnSerIleGluLysSerAspLeuGluGluGluGluGluGluGluGluGluGluGlu 1200
Db 1776 TTGATGGAATCTTGAAACCAAGGGGTAAACAGTAAATCGGATGAGAACTGACTCT 1835
Oy 1201 ValLysLeuLeuLeuGluMetGluLeuLysGluHisIleLeuThrAspSerGluLeuSer 1220
Db 1836 GCTGATATATAGTATCATCTGAGTATATATCAACATCTAGTA---TCTTACAGAAAT 1892
Oy 1221 IleGluLysLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 1240
Db 1893 ATTGAAGACTTCACAAACAAATCAACGCTCTTGTAGTCCCTTGAAGAGCTTGGGAA 1952
Oy 1241 GluMetLysAsnIleThrIleGluArgAsnGluGluGluGluGluGluGluGluGlu 1260
Db 1953 -----ACGAGAGAAAGAGAAACAAACAACT----- 1982
Oy 1261 AlaGluHisAspSerLeuLysGluAspLeuSerGluAsnIleGluGlnSerIleGluThr 1280
Db 1983 -----TCATCCAAATCACTGAGCTTCAAGCTCAAACTTGAGTCCCTTACTGAA 2033
Oy 1281 GlnAspGluLeuArgAlaAlaGluGluGluGluGluGluGluGluGluGluGluGlu 1300
Db 2034 CTAGAACAATCCCGCAATCAGCAGAG-----CATCAATTCGACGCTTGTGATTC 2084
Oy 1301 PheArgGluGln-----LeuLeuAspCysSerValGlyIleSerSer 1314
Db 2085 ATAGTTCGTCAGCGTATATGATACCGTATTTATTTTCAACAAACAGAGATGGCAAT 2144
Oy 1315 ProAsnHisAspAlaValAlaAsnGluGluValSerLeuGluValAsnSerLeu 1334
Db 2145 CCATTACATGCTTCA-----AGCTTACATATGTT---TCTCTT 2180
Oy 1335 GlnSerGluMetLeuArgGluGluArgAspGluLeuGluThrSer----- 1349
Db 2181 GCATCAATCCAAACAGTCAATGATACATCAACAGCTTGTCCCTGCTCCAGTACT 2240
Oy 1350 -----CysLysAlaLeuValSerLysLeuGluLeuLeu 1360
Db 2241 GTTATGTAATCAACAGCGTATAGCGCTAGCGCTGCTTAAACAGTTGCAGAAAT 2300
Oy 1361 ArgAlaHisValLysSerValGluGluAsnLeuGluIleThrLysLysLeuAsnGlu 1380
Db 2301 TTGAGAACTACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 2342
Oy 1381 LeuGluLysGluIleLeuGluLysSerGluGluGluGluGluGluGluGluGluGlu 1400
Db 2343 -----CAAAATGAGCAGCTTGAGAAACTTCACAAACAAACAAACAAACAAACAA 2381
Oy 1401 AsnLeuLysGluAspAsnAsnLysLysGluGluAlaGluGluGluGluGluGluGlu 1420
Db 2382 GATTGTGCGATCCAAATATCCAAATTTCTACCCAGCTA---GATTGTGCTTAAACCT 2438
Oy 1421 AsnGlnPhe-----SerLeuGluGlu 1427
Db 2439 TATGAATCTGCAAGATATGTTGAAGATATCGTCGAGAAATATACATCACTTCAAG 2498
Oy 1428 -----ValPheSerGluGluGluGluGluGluGluGluGluGluGluGluGlu 1443
Db 2499 AGAAATCAGAAACTCACTCCACAACTCAAAAGCAAGAACAGATTCATAGCAGTACT 2558
Oy 1444 AlaGlnLeuLysAlaAlaGluGluGluGluGluGluGluGluGluGluGluGluGlu 1463
Db 2559 CAAAGATTGAGAGCAATATGAGAGCTGCTGCGAGAA----- 2600
Oy 1464 ValGlnThrAlaAsnThrAsnLeuValGluGluLysLeu-----GluThrProLeu 1480
Db 2601 GTTAGAGCAAGAAATTTGAAGAAAGAAAGAAATGCTTAAATTTGCTGAGATGCTCTT 2660
Oy 1481 GlnAlaAspHisGluGluAspSerIleAspArgArgSerGluGluGluGluGluGluVal 1500
Db 2661 TCTCAGCAAAAGAGAGCTTTGTAGCTGAACAAAGGGGCAAAACCTTA----- 2708

```

QY	1845	glnlleserValleuIngInIuysIlleserLeuLeuIngIuaInImerLeuTyAsnVal	1864
Db	3591	gtccgctgacgattttggaaacaaacacacagaaagcaaatccacatttggacagtraa	3550
QY	1865	AlaThr-----ValIysgluThrLeuSerGlu	1873
Db	3651	gcattcttggagagaaacagacagacattttaaAGATGAAAGTTCATGTCGC	3710
QY	1874	ArgAspAspLeuAaInIserIysGlnIhIsLeuPheSerGluIleGluThrLeuSer---	1892
Db	3711	tgtgaagatcttggaaacaaacaaacagatttcttCATGATGATGATGCAAAATTAAGTGAC	3770
QY	1893	-----LeuSerLeu-----	1895
Db	3771	aaGTCGTGGCTCTGTGAAGAAAGGTGTACAAGTCCACTGAATGTATCTCTCAGTGA	3830
QY	1896	-----LysGluIysGlu	1899
Db	3831	GAAGAAATCTCAAGAACAAATTTTGAAATTCAGATTATTACGACGAGAAAAAGAA	3890
QY	1900	-----PheAlaLeuGluIleAlaGlu-----	1906
Db	3891	ATTGCTGAATAGCTTTGAGTGGCTCAGGTGTGAGAGTCTGCCTTTCACAAAGGTT	3950
QY	1907	-----LysAspIysAlaAspAlaAlaArgIysThr	1916
Db	3951	GAACTTTAAAGAAAGACGTGCGAGAACTCGAAGATGACTAAATGCTGAAGGAGAA	4010
QY	1917	IleAspIleThrGluIysIleserAsnIleGluIleLeuLeuSerGlnAlaIleThrAsn	1936
Db	4011	gtccaggaatactccaaacaaacattggctcagcagaaacatgaaagaaatgaaac	4070
QY	1937	LeuIysGluThrLeuTyThrGluIysSerIleGlnIysGluIleAlaIleu	1956
Db	4071	ATGAATGACTATTGAGACCAATAATATGCTTAAGAGAGAGAAAGACGCTGAACAG	4130
QY	1957	AsnThrGlnIhIsleuArgIuThrLeuIysSerIysAspLeuAlaLeuGlyLysMetGlu	1976
Db	4131	GATCTACGCAATGCGACAGAAAGGTGAGAAATCTGAAATGTGATTTTACCCCTTACA	4190
QY	1977	GlnGluArgAspGluAlaAlaAsnIysValIleAlaLeuThrGluIysMetSerLeu	1996
Db	4191	GAAGCAATGCTGAGTGTGAGAAACGGTATGTTTCAGCAGAGAAAGAGCTTTA	4250
QY	1997	GluGluGluIle-----AsnGluAsnValThrThrLeuIysGluGly	2010
Db	4251	GAGAGAGATGTCAAACGTTGGAAGACGTAACACACTCTAGTAAGT-----	4298
QY	2011	GluGlyGluIysGluThrPheTyThrLeuIleArgProSerIysGlnIserSerGln	2030
Db	4299	-----CAACGAAAGATCCAGAT	4316
QY	2031	MetGluIleuArgGluSerLeuIysThrIysAspLeuIleuGluGluIleGluIys	2050
Db	4317	ACAGAAAGATATCGAGACTCTTTCGAAAGAAAGTTCAT-----ACTAAG	4364
QY	2051	GluIleSerGlnAlaThrAsnGluIleIysAsnIleuThrAlaIysIleserSerLeuGlu	2070
Db	4365	CGTATTCACAAATTCGACAGAAATTTGTAGACTTAAAGCTGAATTCGAACATCAAT	4424
QY	2071	GluGluIleLeuGlnAsnIleSerIleLeuAsnGluAlaValSerGluArgIleuAsnLeu	2090
Db	4425	GCATCTTTGACTACACACAGAACTTAAATTCAGAGTCTGAAGAGAACTTAAATTAAGTA	4484
QY	2091	ArgIhIsleuIysGlnIleuValSerGluLeuGluIleuSerLeuThrLeuIysSer	2110
Db	4485	AGAACTGAAGAAAGAACCATTCAGAAAGACTTAAGTGCACAAATATGATATCCAGAA	4544
QY	2111	ArgAspHisAlaPheAlaGlnIserIysArgIuIysAspGluAlaValAsnIysIleAla	2130
Db	4545	AAAGTCAAAACTATTACTCAAGTTAAAGAAATTTGGACGAGGTACAAAGCTCAATATGAA	4604
QY	2131	SerLeu---AlaGluGluIleIysIleuThrIysGluMetCaspGluPheArgAspSer	2149

Db 4605 GAAGTTAAGACACAGAGATAGGTATGAGACATGCGCTCACTCCCTGGAACCAT 4664
 Qy 2150 Lysglu-----SerleuglnlucInserHisleuSerGlulucCysThr 2166
 Db 4665 CAGAGCAGCATGTTTTCCTCAGCAAAATG--CAGGAACCTCAAGAACGCTCAACCA 4721
 Qy 2167 TyrLysThrIleuGlulMetleuLysglInlInLysgluAspIleAsnAsnLysleuLa 2186
 Db 4722 GCTGAACCAAAATCAAAATCTTGAAGTCAAGTACGAATCTGCACAGACATTTACT 4781
 Qy 2187 GluLysValLysgluValAspIleuLysglInHisleuSerSerleuLysglInleu 2206
 Db 4782 GAAAAAGAGACAGAGCAAGAAATCTCCAGAACAGACTGTCACACTTCAGCTCAACTT 4841
 Qy 2207 AspelInleuMetGlulLeuArgAsnGlulLysleuArgAsnTyrGlulucCysgluLys 2226
 Db 4842 TCACGACTTCTCGAGACTTCTCAGATAGAACACACAGAGAGCAGCAGCTC--CGA 4895
 Qy 2227 MetAspIleMetgluLysglu----- 2233
 Db 4896 CAACAGATTAAGTAAAGAAAGAAAAACCAAGAGCTATTGTAGCAGCAAAAGTCAAAA 4955
 Qy 2234 IleserValleuArgleuMetGlnAsnGlulProglInglIngluLysAspValAlaGlu 2253
 Db 4956 ATTCACACTAGCTGAGGTGTAAGATCAGCTAACTAAAGAAATGAGAGAGCTTAAACA 5015
 Qy 2254 ArgmetAspIleleuGlulSerArgAsnGlulInleuGlulMetgluLysIleser 2273
 Db 5016 AGGATGAGAGCTTAGAT--CAGCAGAAAGATGATGGATTGCTTCCGCTTCT 5066
 Qy 2274 AlaValTyrSerGlulInHisThrleuLeuSerSerleuSerSerGlulInleuLysglu 2293
 Db 5067 GCGCTAAAGTCCCAATATGAAGCTGATTTAGTGGCTTGAAGAAAGAACTCAGGAGCAT 5126
 Qy 2294 ThrIleuAlaHisLysHisCysMetleuAsnIleLysglu--SerleuSerSerThrleu 2313
 Db 5127 CAACAGAGAGACAC-----CTTGACAGAGAGATGAGCCCAAGAACTCTCTAT 5174
 Qy 2313 ArgSerPhegluSerleuGlulInHisValLysleuAsnThrIleuGlulInleu 2333
 Db 5175 AAGGTCCCTGACACAGCAGACAGATCAGATTGAAA--CAACTCCACACTCT 5225
 Qy 2333 ulLeuAsnLysPheLysValValTyrArgThrAlaValLysgluAsnHisSerleu 2353
 Db 5226 G-----CTGAAGAG-- 5235
 Qy 2353 elYsAspTyrGlulLysAspLeuAlaGluLulLysArgHisAspGlulLeuArgleu 2373
 Db 5236 -----GAAATGCCAGCAGATCAGACCCACCAACAGCAATATCAAGCCA 5279
 Qy 2373 n-----leuGlucylserleuGlulInHis-----GlyArgLysTr 2384
 Db 5280 ACTCTCTGTTGTTACTCCAAAGT-AAAGTACAGCTGCACCTATGGTGGAAATAATGTC 5338
 Qy 2384 pSerAspSerAlaSerGlulInleuLysPheCysglulIleuPheLeuAsnGlulLeu 2404
 Db 5339 AACACCCAGGCTAGTATCCGCCCAATGGTTACACCTGCACTGTACAAATCCACATAC 5398
 Qy 2404 uPheLysAlaAsnIleleuSerValGlnAspAspPheSerGlulValGlnValPh 2424
 Db 5399 TACCCCAACAGCTACATGATGATGACCTACACAAAGTGAA----- 5438
 Qy 2424 eleuAsnGlnValGlySerThrleuGlulInglulInleuLysLysgluPheMetG 2444
 Db 5439 -----TCACAGAAAGCTATGACAGTCAAGAGGCGCTGTGCA 5473
 Qy 2444 nTrpLeuGlulInleuPheLysAspLeuHisValAlaLysLysleuSerGlulLysMetG 2464
 Db 5474 ACATGTTCCATTTTGGAAACACAAAGTATCCGTTCTCTACTAGTCTTAATGTCCA 5533
 Qy 2464 nGlnGluAsnArgArgIleAlaSerThrIleGlnleuLeuThrLysArgleuLysAlaVal 2484
 Db 2484 ----- 2484

Db 5534 GCGTTCTATCTCTCAACCTATTTTAACTGTTTCAGCAACAAACA--CAGGCTACAGCTTT 5590
 Qy 2484 lValGlnSerLysIleGlnArgGluIleThrValLysleuAsnGlnPheGluAlaLysle 2504
 Db 5591 TGTGCAACCCACTCAACAGAGTCAT-----CCTCAGATTGACCTGCGCA 5635
 Qy 2504 uGlnGluLysLysgluGlnAsnLysglulLeuMetArgArgMetGlnHisIleGlyProse 2524
 Db 5636 TCAGAGTTATGTTCAACATATGTAGAGGTTTTCAGAGTTTCCAGATTGACAGGCGCTTC 5695
 Qy 2524 ValSerValMetgluGlulInleuAsnAlaArgleuLysglulLeuLysThrValGlnAs 2544
 Db 5696 TACTTCC-----ACAGCACTATTGTCACAGTTTGCGGTACCCCGAG 5737
 Qy 2544 pGluSerLysLysleuGlnSerArgIleLysMetleuGluAsnGluLysleuValLys 2564
 Db 5738 TTCTCTTGGCCCAAGCGTACAGT-----GAGAGGAA----- 5771
 Qy 2564 sAspAspAlaMetHisLysglulLysValAlaIleleuGlnAspLysleuSerArg 2584
 Db 5772 -----GAGATAGCCATCGAAGC 5791
 Qy 2584 gAsnAlaGluAlaGluLeuAsnAlaMetGlnValLysleuThrLysgluAspAsnle 2604
 Db 5792 ATCAGACCAAGTCTGTGATGATACAGTGAATGGCTCTCCAAAGAG----- 5840
 Qy 2604 uGlnAlaAlaMetLysglulIleGluAsnleuGlnLysMetValAlaLysgluValAlaPr 2624
 Db 5841 -----TTGAAAGTGTCAACCC 5857
 Qy 2624 cTyrLysgluGluIleAspAsnLeuLysThrLysValLysIleGluMetGluLys 2644
 Db 5858 TGTAGCACTGAGAAAGAAAGTATATGCAAGAAAGTACTGATGGAGAGTGAACATCA 5917
 Qy 2644 elYsTyrSerLysAlaThrAspGlnGluLea-----Tyrle 2657
 Db 5918 GGTATACAAACCGAGATTCTCAAGATTCTGAGAGAGAGTATCCACGAGATTAAT-- 5975
 Qy 2657 ulYsSerCysleuGluAspLysglulGluLysArgArgleuLysglulLeuArgAr 2677
 Db 5976 -----ACACCTATGAGAGACGTGAAGAAACCTCTCAGTCTCTCAAAATGATCTTGGCC 6031
 Qy 2677 gAlaGlnAlaAspAsnAspThrThr----- 2685
 Db 6032 ACTCAATCAGATCAGCAGACAGCAACTTCATCCAGATGTCAAGCAAGAGATGA 6091
 Qy 2686 ----ValCysValProLysAspTyrGlnLysAla----- 2695
 Db 6092 TGTCAATTGTAATTGACAGTATGATGAAGAAAGAGATGAGAAATGATGATGATGA 6151
 Qy 2695 ----- 2695
 Db 6152 AGATGACACAGGATGGAGATGAGGTGAGATAGTAATGAAGAACTGTATGCCGA 6211
 Qy 2696 -----SerThrPheProValTh 2701
 Db 6212 TGGCAATGATGTTATGAAGCTGATGATGAGGTGAGGTGATGAGACTGATCAGAGTAC 6271
 Qy 2701 r-----CysGlyGlyLysSerGlyTlleValGlnSerThrAlaMetleuVa 2716
 Db 6272 AGAAACAGAGAAAGTAATGAGGTGAGGTGAAGGT----- 6305
 Qy 2716 lLeuGlnSerGluLysAlaAlaLeuGlnArgIleuSerHisTyrLysLysTyrH 2736
 Db 6306 -----AATCAGAGAGCTGCTGATTCAAAACAGTGTGAAGAAATACAGGTGCTGC 6358
 Qy 2736 sHisLeuSerArgThrMetSerSerSerGluAspArgLysThrLysAlaLysSerAs 2756
 Db 6359 AGAATCTCTTTTCTCAGAGAGTTCATGAGAAACAAACAGCATATCAGCA--TCTGA 6415
 Qy 2756 pAlaHisSerSerHisThrGlySerSerHisArgGlySerProHisLysThrGluThrTy 2776
 Db 6416 AAGACAGGCCCTCGAGCACTCACTCAACGAGAGCCACACAT----- 6461

Search completed: May 23, 2003, 19:43:25
Job time : 1712 secs.

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 01:00:29 ; Search time 473 Seconds
(without alignments)
6241.164 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626

Sequence: 1 gattccgagtcgtagtag.....ttaaaaaaaaaacggaattc 9626

Scoring table: OLIGO_NDC
Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued_Patents_NA:*

- 1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
- 2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
- 3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
- 4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
- 5: /cgn2_6/prodata/2/lna/PCUTS.COMB.seq:*
- 6: /cgn2_6/prodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	0.2	10564	1 US-08-206-176-5	Sequence 5, Appl1
2	19	0.2	423	3 US-08-906-769-134	Sequence 134, App
3	19	0.2	423	3 US-08-906-616-134	Sequence 134, App
4	19	0.2	423	3 US-08-639-075A-134	Sequence 134, App
5	19	0.2	423	4 US-09-012-431-134	Sequence 134, App
6	19	0.2	423	4 US-09-012-692-134	Sequence 134, App
7	19	0.2	423	4 US-08-906-613-134	Sequence 134, App
8	19	0.2	445	1 US-08-207-904-14	Sequence 14, Appl
9	19	0.2	658	4 US-09-328-111-665	Sequence 665, App
10	19	0.2	4080	1 US-08-375-300-1	Sequence 1, Appl1
11	19	0.2	4080	3 US-09-177-431-1	Sequence 1, Appl1
12	19	0.2	4080	5 US-09-1595-16930-1	Sequence 1, Appl1
13	19	0.2	21338	4 US-08-961-527-20	Sequence 20, Appl
14	19	0.2	42571	4 US-09-810-347-3	Sequence 3, Appl1
15	18	0.2	31	2 US-08-859-998-170	Sequence 170, App
16	18	0.2	31	4 US-09-225-928-170	Sequence 170, App
17	18	0.2	147	4 US-09-269-833D-20	Sequence 20, Appl
18	18	0.2	315	2 US-08-743-200-1	Sequence 1, Appl1
19	18	0.2	432	2 US-09-001-944-11	Sequence 11, Appl
20	18	0.2	432	4 US-09-240-004A-11	Sequence 11, Appl
21	18	0.2	435	2 US-09-001-944-9	Sequence 9, Appl1
22	18	0.2	435	4 US-09-240-004A-9	Sequence 9, Appl1
23	18	0.2	435	4 US-09-269-833D-30	Sequence 30, Appl
24	18	0.2	441	4 US-09-269-833D-28	Sequence 28, Appl
25	18	0.2	450	4 US-09-240-004A-13	Sequence 13, Appl
26	18	0.2	453	4 US-09-269-833D-29	Sequence 29, Appl
27	18	0.2	457	4 US-09-269-833D-26	Sequence 26, Appl

28	18	0.2	498	2 US-09-001-944-1	Sequence 1, Appl1
29	18	0.2	498	2 US-09-001-944-3	Sequence 3, Appl1
30	18	0.2	498	2 US-09-001-944-5	Sequence 5, Appl1
31	18	0.2	498	2 US-09-001-944-7	Sequence 7, Appl1
32	18	0.2	498	4 US-09-240-004A-1	Sequence 1, Appl1
33	18	0.2	498	4 US-09-240-004A-3	Sequence 3, Appl1
34	18	0.2	498	4 US-09-240-004A-5	Sequence 5, Appl1
35	18	0.2	498	4 US-09-240-004A-7	Sequence 7, Appl1
36	18	0.2	501	4 US-09-269-833D-3	Sequence 3, Appl1
37	18	0.2	501	4 US-09-269-833D-27	Sequence 27, Appl
38	18	0.2	517	4 US-09-269-833D-13	Sequence 13, Appl
39	18	0.2	517	4 US-09-269-833D-23	Sequence 23, Appl
40	18	0.2	906	4 US-08-976-259-98	Sequence 98, Appl
41	18	0.2	955	4 US-09-641-638-9	Sequence 9, Appl1
42	18	0.2	955	4 US-09-641-638-10	Sequence 10, Appl
43	18	0.2	955	4 US-09-641-638-11	Sequence 11, Appl
44	18	0.2	959	4 US-09-724-510-1	Sequence 1, Appl1
45	18	0.2	959	4 US-09-723-216-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-206-176-5
; Sequence 5, Application US/08206176
; Patent No. 5639940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Dalrymple, Michael A
; APPLICANT: Prunkard, Donna E
; APPLICANT: Foster, Donald C
; TITLE OR INVENTION: Production of Fibrinogen in Transgenic
; TITLE OF INVENTION: Animals
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: human fibrinogen gamma chain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
; ..2603, 4211..4341, 4645..4778, 5758..5942, 7426
; ..7703, 9342..9571)
; LOCATION: ..7703, 9342..9571)
; US-08-206-176-5

Query Match 0.2%; Score 21; DB 1; Length 10564;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2975 TTCCTCAAGACGCTTGAGAAA 2995
|||||
Db 3246 TTCCTCAAGACGCTTGAGAAA 3266

RESULT 2

US-08-906-769-134
; Sequence 134, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..423
US-08-906-769-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGAGA 1842
|||||
Db 222 AACTACAGTCAAGAGAGA 240

RESULT 3

US-08-906-616-134
; Sequence 134, Application US/08906616
; Patent No. 6121035
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..423
US-08-906-616-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGAGA 1842
|||||
Db 222 AACTACAGTCAAGAGAGA 240

RESULT 4

US-08-639-075A-134
; Sequence 134, Application US/08639075A
; Patent No. 6150125
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,075A
FILING DATE: 24-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-08-639-075A-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 5
US-09-012-431-134
Sequence 134, Application US/09012431
Patent No. 6180383
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,431
FILING DATE: 23-Jan-1998

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-09-012-431-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 6
US-09-012-692-134
Sequence 134, Application US/09012692
Patent No. 6214579
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,692
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-09-012-692-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 7
US-08-906-613-134
Sequence 134, Application US/08906613
Patent No. 6232096
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423

US-08-906-613-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGGA 1842
DB 222 AACTACAGTCAAGAGGA 240

RESULT 8
US-08-207-904-14
Sequence 14, Application US/08207904
Patent No. 5477002

GENERAL INFORMATION:
APPLICANT: Tuttle, AnnMarie
APPLICANT: Crossland, Lyle D.
TITLE OF INVENTION: Another-Specific CDNA Sequences, Genomic
TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/207,904
FILING DATE:
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/908,242
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.

REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: CGC 1624
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Nicotiana tabacum
INDIVIDUAL ISOLATE: Ant68

FEATURE:
NAME/KEY: CDS
LOCATION: 2..445
US-08-207-904-14

Query Match 0.2%; Score 19; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4924 CAATGTCGAATCACTCAA 4942
DB 301 CAATGTCGAATCACTCAA 319

RESULT 9
US-09-328-111-665
Sequence 665, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adrian
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
EARLIER FILING DATE: 1998-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 665
LENGTH: 658
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(658)
OTHER INFORMATION: n - A,T,C or G
US-09-328-111-665

Query Match
Best Local Similarity 100.0%; Pred. No. 69;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3813 TGCAGCTGGAGATCTGCA 3831
DB 24 TGCAGCTGGAGATCTGCA 42

RESULT 10
US-08-375-300-1
Sequence 1, Application US/08375300
Patent No. 5679566
GENERAL INFORMATION:
APPLICANT: Feng, He
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. P.
REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 04020/046001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-375-300-1

Query Match
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4166 GAAGAGATGAGCTTCAAA 4184
DB 862 GAAGAGATGAGCTTCAAA 880

RESULT 11
US-09-177-431-1
Sequence 1, Application US/09177431
Patent No. 6071700
GENERAL INFORMATION:
APPLICANT: He, Feng
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-177-431-1

Query Match
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4166 GAAGAGATGAGCTTCAAA 4184

Db 862 GAAAGAGATGAGCTTCAA 880

RESULT 12

PCT-US95-16930-1

Sequence 1, Application PC/TUS9516930

GENERAL INFORMATION:

APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL

TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE

TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF

TITLE OF INVENTION: NONSENSE-MEDIATED MRNA DECAY

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street Suite 3100

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/16930

FILING DATE: 27-DEC-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/375,300

FILING DATE: 20-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Fasse, J. Peter

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 04020/046WO1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)542-5070

TELEFAX: (617)542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4080 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

PCT-US95-16930-1

Query Match 0.2%; Score 19; DB 5; Length 4080;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4166 GAAAGAGATGAGCTTCAA 4184

Db 862 GAAAGAGATGAGCTTCAA 880

RESULT 13

US-08-961-527-20

Sequence 20, Application US/08961527

Patent No. 6420135

GENERAL INFORMATION:

APPLICANT: Charles Kunach

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 Inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 21338 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-20

Query Match 0.2%; Score 19; DB 4; Length 21338;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3446 AAAAAATCTTTTTCAG 3464

Db 1768 AAAAAATCTTTTTCAG 1786

RESULT 14

US-09-810-347-3

Sequence 3, Application US/09810347

Patent No. 6461847

GENERAL INFORMATION:

APPLICANT: Ye, Jane et al.

TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES

FILE REFERENCE: C1001169

CURRENT APPLICATION NUMBER: US/09/810,347

CURRENT FILING DATE: 2001-03-19

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 42571

TYPE: DNA

ORGANISM: Human

US-09-810-347-3

Query Match 0.2%; Score 19; DB 4; Length 42571;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3109 GATTAATGTTATTAAAGCA 3127

Db 39653 GATTAATGTTATTAAAGCA 39671

RESULT 15

US-08-859-998-170

Sequence 170, Application US/08859998

Patent No. 5994076

GENERAL INFORMATION:

APPLICANT: Chenchik, Alex

APPLICANT: Johadze, George

APPLICANT: Biblashvili, Robert

```

: TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
: TITLE OF INVENTION: EXPRESSION
: NUMBER OF SEQUENCES: 1375
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson, P.C.
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: CA
: COUNTRY: US
: ZIP: 94025
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/859,998
: FILING DATE: 21-MAY-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Field, Bret E.
: REGISTRATION NUMBER: 37,620
: REFERENCE/DOCKET NUMBER: 09096/002001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-322-5070
: TELEFAX: 415-854-0875
: INFORMATION FOR SEQ ID NO: 170:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: FEATURE:
: OTHER INFORMATION: oligonucleotide primer
: US-08-859-998-170
:
Query Match 0.28; Score 18; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5338 ATGCTGCTAATGAGAA 5355
DB 7 ATGCTGCTAATGAGAA 24

```

Search completed: May 23, 2003, 08:05:39
 Job time : 910 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 07:32:30 ; Search time 1097 Seconds
(without alignments)
11586.846 Million cell updates/sec

Title: US-09-150-867-2
Perfect score: 9626
Sequence: 1 gaattccggagtcgtagg.....ttaaaaaaaaaaaggaattc 9626

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 828747 seqs, 660231138 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCr_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24	0.2	1100	10 US-09-263-959-276	Sequence 276, App
C 2	24	0.2	1277	10 US-09-263-959-269	Sequence 269, App
C 3	24	0.2	640681	10 US-09-790-988-1	Sequence 1, Appl1
C 4	24	0.2	684973	10 US-09-263-959-1	Sequence 1, Appl1
C 5	23	0.2	7684	10 US-09-070-927A-16	Sequence 16, Appl1
C 6	20	0.2	1215	9 US-09-925-299-79	Sequence 9, Appl1
C 7	20	0.2	1215	10 US-09-925-299-79	Sequence 79, Appl1
C 8	20	0.2	4510	12 US-10-005-983-1	Sequence 1, Appl1
C 9	20	0.2	7992	9 US-09-893-519A-140	Sequence 140, App
C 10	20	0.2	8493	9 US-10-071-766-51	Sequence 51, Appl1
C 11	20	0.2	180557	12 US-10-003-806-6	Sequence 6, Appl1
C 12	20	0.2	180557	12 US-10-003-806-6	Sequence 6, Appl1
C 13	20	0.2	402850	9 US-09-844-653-5	Sequence 5, Appl1
C 14	19	0.2	70	10 US-09-815-242-371	Sequence 371, App
C 15	19	0.2	231	9 US-09-864-761-31477	Sequence 31477, A
C 16	19	0.2	247	9 US-09-918-995-18268	Sequence 18268, A
C 17	19	0.2	344	9 US-09-918-995-29936	Sequence 29936, A
C 18	19	0.2	348	10 US-09-864-761-14056	Sequence 14056, A
C 19	19	0.2	355	10 US-09-864-761-28258	Sequence 28258, A

20	19	0.2	358	10 US-09-960-352-13853	Sequence 13853, A
C 21	19	0.2	361	10 US-09-864-761-22536	Sequence 22536, A
C 22	19	0.2	364	10 US-09-783-590-5629	Sequence 5629, App
C 23	19	0.2	388	9 US-09-918-995-3711	Sequence 3711, App
C 24	19	0.2	430	9 US-09-918-995-34048	Sequence 34048, A
C 25	19	0.2	432	9 US-09-918-995-34983	Sequence 34983, A
C 26	19	0.2	440	10 US-09-960-352-9139	Sequence 9139, App
C 27	19	0.2	443	10 US-09-764-877-900	Sequence 900, App
C 28	19	0.2	455	10 US-09-864-761-11679	Sequence 11679, A
C 29	19	0.2	457	10 US-09-864-761-5775	Sequence 5775, App
C 30	19	0.2	577	10 US-09-864-761-14948	Sequence 14948, A
C 31	19	0.2	591	10 US-09-864-761-7654	Sequence 7654, App
C 32	19	0.2	639	9 US-10-066-543-666	Sequence 666, App
C 33	19	0.2	658	10 US-09-879-536-665	Sequence 665, App
C 34	19	0.2	740	10 US-09-910-943-690	Sequence 690, App
C 35	19	0.2	771	12 US-10-001-843-64	Sequence 64, Appl1
C 36	19	0.2	1100	10 US-09-263-959-275	Sequence 275, App
C 37	19	0.2	1160	9 US-09-764-891-10183	Sequence 10183, A
C 38	19	0.2	1203	10 US-09-263-959-239	Sequence 239, App
C 39	19	0.2	1381	9 US-10-023-282-82	Sequence 82, Appl1
C 40	19	0.2	1439	10 US-09-882-830A-51	Sequence 51, Appl1
C 41	19	0.2	1728	9 US-09-938-842A-2550	Sequence 2550, App
C 42	19	0.2	1838	9 US-09-938-842A-4073	Sequence 4073, App
C 43	19	0.2	1880	10 US-09-731-872-78	Sequence 78, Appl1
C 44	19	0.2	2000	9 US-09-938-842A-4182	Sequence 4182, App
45	19	0.2	2109	10 US-09-070-927A-488	Sequence 488, App

ALIGNMENTS

RESULT 1

US-09-263-959-276/C
Sequence 276, Application US/09263959
Patent No. US20020150891A1

GENERAL INFORMATION:

APPLICANT: Hoood, Leroy E.

APPLICANT: Rowen, Lee

APPLICANT: Koop, Ben F.

TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U

NUMBER OF SEQUENCES: 1279

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/263, 959

FILING DATE: 05-MAR-1999

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: McMaisters, David D.

REGISTRATION NUMBER: 33,963

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 276:

SEQUENCE CHARACTERISTICS:

LENGTH: 1100 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-263-959-276

Query Match 0.2%; Score 24; DB 10; Length 1100;

Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAGCAAAAGCTGATGCT 5878
|||||
DB 689 GCAGAGAGCAAAAGCTGATGCT 666

RESULT 2

US-09-263-959-269/c
; Sequence 269, Application US/09263959
; Patent No. US20020150891A1
; GENERAL INFORMATION:
; APPLICANT: Hood, Leroy E.
; APPLICANT: Rowen, Lee
; APPLICANT: Koop, Ben F.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI
; NUMBER OF SEQUENCES: 1279
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,959
; FILING DATE: 05-MAR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 269:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-263-959-269

Query Match 0.2%; Score 24; DB 10; Length 1277;

Best Local Similarity 100.0%; Pred. No. 0.57;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAGCAAAAGCTGATGCT 5878
|||||
DB 758 GCAGAGAGCAAAAGCTGATGCT 735

RESULT 3

US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: MATANABE, HIDEMI
; APPLICANT: HATORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHITOKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 640681

; TYPE: DNA

; ORGANISM: Buchnera sp.

US-09-790-988-1

Query Match 0.2%; Score 24; DB 10; Length 640681;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3957 AAAACATTGAGCAGTCATTTGAA 3980
|||||
DB 505797 AAAACATTGAGCAGTCATTTGAA 505820

RESULT 4

US-09-263-959-1/c
; Sequence 1, Application US/09263959
; Patent No. US20020150891A1
; GENERAL INFORMATION:
; APPLICANT: Hood, Leroy E.
; APPLICANT: Rowen, Lee
; APPLICANT: Koop, Ben F.
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U
; NUMBER OF SEQUENCES: 1279
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,959
; FILING DATE: 05-MAR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 920010.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 684973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-263-959-1

Query Match 0.2%; Score 24; DB 10; Length 684973;

Best Local Similarity 100.0%; Pred. No. 0.94;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAGCAAAAGCTGATGCT 5878
|||||
DB 163714 GCAGAGAGCAAAAGCTGATGCT 163691

RESULT 5

US-09-070-927A-16
; Sequence 16, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch

Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 7684 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-070-927A-16
Query Match 0.2%; Score 23; DB 10; Length 7684;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3973 AATTGAACACAGATGAATTAA 3995
|||||
DB 4287 AATTGAACACAGATGAATTAA 4309
RESULT 6
US-09-925-299-79
Sequence 79, Application US/09925299
Publication No. US20030040617A9
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 79
LENGTH: 1215
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-299-79

Query Match 0.2%; Score 20; DB 9; Length 1215;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9599 ATTTAAATTAAAAAAA 9618
|||||
DB 1178 ATTTAAATTAAAAAAA 1197
RESULT 7
US-09-925-299-79
Sequence 79, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 79
LENGTH: 1215
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-299-79
Query Match 0.2%; Score 20; DB 10; Length 1215;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9599 ATTTAAATTAAAAAAA 9618
|||||
DB 1178 ATTTAAATTAAAAAAA 1197
RESULT 8
US-10-005-983-1
Sequence 1, Application US/10005983
Patent No. US20020116730A1
GENERAL INFORMATION:
APPLICANT: Allen, Keith D.
TITLE OF INVENTION: TRANSGENIC MICE CONTAINING PERK PROTEIN
FILE REFERENCE: R-517
CURRENT APPLICATION NUMBER: US/10/005,983
CURRENT FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/246,676
PRIOR FILING DATE: 2000-11-07
PRIOR APPLICATION NUMBER: US 60/311,018
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/324,765
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 60/326,148
PRIOR FILING DATE: 2001-09-28
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 4510
TYPE: DNA
ORGANISM: Mus musculus
US-10-005-983-1
Query Match 0.2%; Score 20; DB 12; Length 4510;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9599 ATTTAAATTAAAAAAA 9618

DB 4476 ATTATTAAGAAAAA 4495

RESULT 9
US-09-893-519A-140

; Sequence 140, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BUTTMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESTIVA, Thamar
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MCCOY, Daniel
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAO, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTITUMOR COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/16348-US2
; CURRENT FILING DATE: 2001-06-28
; PRIOR FILING DATE: 2001-06-28
; PRIOR FILING DATE: 2000-06-29
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/Z15005.1
; DATABASE ENTRY DATE: 1993-01-10
; RELEVANT RESIDUES: (1)..(7992)
US-09-893-519A-140

Query Match 0.2%; Score 20; DB 9; Length 7992;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8012 CCATTAAGAAAGAAATTGA 8031
DB 7240 CCATTAAGAAAGAAATTGA 7259

RESULT 10
US-10-071-766-51
; Sequence 51, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huet-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 51
; LENGTH: 8493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 441283.3

US-10-071-766-51

Query Match 0.2%; Score 20; DB 9; Length 8493;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8012 CCATTAAGAAAGAAATTGA 8031
DB 7330 CCATTAAGAAAGAAATTGA 7349

RESULT 11
US-10-003-806-6/c

; Sequence 6, Application US/10003806
; Patent No. US20020119929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 180557
; TYPE: DNA
; ORGANISM: Human
US-10-003-806-6

Query Match 0.2%; Score 20; DB 12; Length 180557;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5449 AAAAGATCTGAACAAAAG 5468
DB 91760 AAAAGATCTGAACAAAAG 91741

RESULT 12
US-10-003-806-9/c

; Sequence 9, Application US/10003806
; Patent No. US20020119929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 180557
; TYPE: DNA
; ORGANISM: Human
US-10-003-806-9

Query Match 0.2%; Score 20; DB 12; Length 180557;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5449 AAAAGATCTGAACAAAAG 5468
DB 91760 AAAAGATCTGAACAAAAG 91741

```
RESULT 13
US-09-844-653-5
; Sequence 5, Application US/09844653
; Publication NO. US20030054347A1
; GENERAL INFORMATION:
; APPLICANT: Richards, Julia
; APPLICANT: Rozsa, Frank
; TITLE OF INVENTION: Detecting and Treating Eye Disease
; FILE REFERENCE: UM-06105
; CURRENT APPLICATION NUMBER: US/09/844,653
; CURRENT FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patent version 3.0
; SEQ ID NO 5
; LENGTH: 402850
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (46565)..(46565)
; OTHER INFORMATION: The n at this position can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (117443)..(118256)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (118272)..(118736)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (120172)..(120640)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (122654)..(122654)
; OTHER INFORMATION: The n at this position can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (203527)..(203963)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (215340)..(215340)
; OTHER INFORMATION: The n at this position can be a, c, t, or g.
; NAME/KEY: misc-feature
; LOCATION: (222415)..(222416)
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
; OTHER INFORMATION: The n at these positions can be a, c, t, or g.
US-09-844-653-5

Query Match          0.2%; Score 20; DB 9; Length 402850;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1834 AAAGGAGGAGGAGGAGGAGG 1853
Db 300865 AAAGGAGGAGGAGGAGGAGG 300884

RESULT 14
US-09-815-242-371
; Sequence 371, Application US/09815242
; Patent NO. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: E117RA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 371
; LENGTH: 70
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(70)
; OTHER INFORMATION: n - A,T,C or G
US-09-815-242-371

Query Match          0.2%; Score 19; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 1,56+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1819 GCACACACTACAGTCAAG 1837
Db 51 GCACACACTACAGTCAAG 69

RESULT 15
US-09-864-761-31477
; Sequence 31477, Application US/09864761
; Patent NO. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```

Sun May 25 14:36:33 2003

us-09-150-867-2.01.rmpb

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31477
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AB016897.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: NT HIT: AB011399.1, EVALU 1.00e-127
; OTHER INFORMATION: EST HUMAN HIT: BF528927.1, EVALU 2.10e-01
; OTHER INFORMATION: SWISSPROT HIT: Q59146, EVALU 3.60e-01
; US-09-864-761-31477

Query Match          0.2%; Score 19; DB 10; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1876 CAAGATGCGAGAACTGAA 1894
   |||
Db 176 CAAGATGCGAGAACTGAA 194

Search completed: May 23, 2003, 18:39:49
Job time : 14864 secs
```

GenCore version 5.1.5
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2003, 22:45:04 ; Search time 471 Seconds
(without alignments)
6267.666 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626
Sequence: 1 gaattccggagtcgagatag.....ttaaaaaaaaaaggaattc 9626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/p/ptodata/2/1na/5A_COMB.seq:*
- 2: /cgn2_6/p/ptodata/2/1na/5B_COMB.seq:*
- 3: /cgn2_6/p/ptodata/2/1na/6A_COMB.seq:*
- 4: /cgn2_6/p/ptodata/2/1na/6B_COMB.seq:*
- 5: /cgn2_6/p/ptodata/2/1na/PCrUS_COMB.seq:*
- 6: /cgn2_6/p/ptodata/2/1na/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147.4	1.5	3837	US-09-724-517-1	Sequence 1, Appl1
2	147.4	1.5	3837	US-09-641-807A-1	Sequence 1, Appl1
3	147.4	1.5	3837	US-09-723-096-1	Sequence 1, Appl1
4	140	1.5	1023	US-09-724-517-3	Sequence 3, Appl1
5	140	1.5	1023	US-09-641-807A-3	Sequence 3, Appl1
6	140	1.5	1023	US-09-723-096-3	Sequence 3, Appl1
7	135.2	1.4	1421	US-09-592-054-5	Sequence 3, Appl1
8	135.2	1.4	1569	US-09-592-054-3	Sequence 3, Appl1
9	127.2	1.3	4127	US-09-592-054-7	Sequence 3, Appl1
10	116	1.2	4308	US-09-592-054-1	Sequence 1, Appl1
11	98	1.0	1011	US-09-641-806-3	Sequence 3, Appl1
12	98	1.0	1026	US-09-641-806-1	Sequence 1, Appl1
13	95.8	1.0	1542	US-09-724-519-5	Sequence 5, Appl1
14	95.8	1.0	1542	US-09-592-037-5	Sequence 5, Appl1
15	95.8	1.0	1728	US-09-724-519-7	Sequence 7, Appl1
16	95.8	1.0	1728	US-09-592-037-7	Sequence 7, Appl1
17	95.8	1.0	3741	US-09-541-782-9	Sequence 9, Appl1
18	95.8	1.0	3741	US-09-723-820-9	Sequence 9, Appl1
19	90	0.9	29793	US-09-302-812-38	Sequence 38, Appl1
20	90	0.9	29793	US-09-511-477-38	Sequence 38, Appl1
21	90	0.9	29793	US-09-511-507-38	Sequence 38, Appl1
22	88	0.9	5093	US-08-468-036-23	Sequence 23, Appl1
23	88	0.9	5093	US-08-376-843-23	Sequence 23, Appl1
24	86.2	0.9	1107	US-09-724-519-9	Sequence 9, Appl1
25	86.2	0.9	1107	US-09-592-037-9	Sequence 9, Appl1
26	85.2	0.9	7218	US-08-232-463-14	Sequence 14, Appl1
27	82.8	0.9	1077	US-09-722-139-3	Sequence 3, Appl1

28	82.8	0.9	1077	US-09-721-832-3	Sequence 3, Appl1
29	82.8	0.9	1077	US-09-721-689-3	Sequence 3, Appl1
30	82.8	0.9	4176	US-09-722-139-1	Sequence 1, Appl1
31	82.8	0.9	4176	US-09-721-832-1	Sequence 1, Appl1
32	82.8	0.9	4176	US-09-721-689-1	Sequence 1, Appl1
33	82	0.9	1230	US-09-572-191-5	Sequence 5, Appl1
34	82	0.9	1230	US-09-723-262-5	Sequence 5, Appl1
35	82	0.9	1230	US-09-723-219-5	Sequence 5, Appl1
36	82	0.9	4757	US-09-572-191-1	Sequence 1, Appl1
37	82	0.9	4757	US-09-723-262-1	Sequence 1, Appl1
38	82	0.9	4757	US-09-723-219-1	Sequence 1, Appl1
39	79.8	0.8	2389	US-08-480-552-18	Sequence 18, Appl1
40	79.8	0.8	2389	US-08-486-382-4	Sequence 4, Appl1
41	79.8	0.8	2389	US-09-235-546-4	Sequence 4, Appl1
42	79.8	0.8	2389	US-08-923-208-18	Sequence 18, Appl1
43	79.8	0.8	2389	US-09-568-315-18	Sequence 18, Appl1
44	79.2	0.8	1149	US-09-724-519-3	Sequence 3, Appl1
45	79.2	0.8	1149	US-09-592-037-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-724-517-1
; Sequence 1, Application US/09724517
; Patent No. 6379941
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: No. 6379941el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/724,517
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3837
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (126)...(1337)
; OTHER INFORMATION: n = a, t, c, or g
US-09-724-517-1

Query Match 1.5%; Score 147.4; DB 4; Length 3837;
Best Local Similarity 51.3%; Pred No. 1.6e-26;
Matches 467; Conservative 0; Mismatches 411; Indels 33; Gaps 4;

QY	275	TTCAATTTCGATCGGTATTAATCTCAAGATCAACAGTCAATTTACAGAGATA	334
DB	136	TTCACTTTGTTGTTGTTTGGCAAAATTCACATCAAGATGAAGTTTAACACATGT	195
QY	335	GCAGTACATCATATACATGACCTTTCGAGGATATATGACACATTTTGCATACGA	394
DB	196	ATTAAGCCCCCTGCTGTCACACTCAATGAGGCTAATGACACCTTTTCCCTATGGA	255
QY	395	GAGCATCTTCAGGCAAGCATACATGATGAGGAACCAATTCATT	444
DB	256	CAAACTGATCTCGGAGACATACACATTTGGAGGGCCATTTCTTCAGTTGGAG	315
QY	445	-----GGGATATATACCCCAAGCATACAGAGATTATTAATTTATTCAGAGATA	496
DB	316	GGCCAAAGAGTATATCTTCCTCGAGCTATTTCAAGAAATTTTCAAGACATCTGMAAT	375
QY	497	CCGACACAGAGATTTCTTTAAGAGTTTCTTTATGAGATTTTACATTAACCTGTGAA	556
DB	376	CCTAGCATTTGACTTATATGTAAGATATCTTATATGAGATGTACAGAGACATCAAGA	435

QY 557 GACCTACTGTGTGATGACAGAGAAAGAGCCCTGGAAATTCGAGGATTTTAATAGA 616
 Db 436 GATCTTTAGATTTGGAGACATTCATGAGATCTTCATCCGAGAAAGTAAAGAA 495
 QY 617 AACGTGATGTGTGACCTGACCTGAGAAACTTGTAAATGTTCTCGAACAATGTAACAG 676
 Db 496 AACCACTGATGTTGGGGCCCAAGGAATGCCATGTGAGAGTGAAGTGAATGAT 555
 QY 677 TGGATCAAAAAAGGTGAAAAAAGACATTTATGAGAGATTAATGATCATAGT 736
 Db 556 CTTTGGAGATGGGGAAATGCAGCAACATACAGTACCTCAATGAATGAGACATCC 615
 QY 737 AGTCGTTCACTACATATTTTGAATGATGTTGAAAGCCGAGAGAAAT-----GAT 790
 Db 616 AGCAGATCAATGCAATTTTACATCAGCAATTTGTCAGTTCAATTAATATGAGAGCA 675
 QY 791 CCCCAATTCAGAGAACTGTGATGAGAGCTGTCAATGGATCTCATTGATTTGATAGT 850
 Db 676 GCTGAAAGTGGATATGATGATTTCCCTCGGCATATTTGCTCAAGTTCCATTGTGAT 735
 QY 851 CTTTGTGCAAGTGAAGAGCAAGCCAACTGAGCTGAGAGTGTGAGACTTAAGAAAGGC 910
 Db 736 TTGGCAGAGATCAGAAAGATTAACCAAAACGGGAATAGTGTGACGTTCAAGAAATCC 795
 QY 911 TGCACATCAACCGCAGCTTGTATTCCTTGACAGGTTATTAAGAAAGCTTAGAGAGC-- 968
 Db 796 ATTCAAATCAATAGTATGATGCTGCTTTAGAGAAATGTAATAGCGCTTGGAGACCA 855
 QY 969 -GCCAGCTGTGATTTATTAATCAAGAGCAAACTCAGCAATTTCTCCAAAT 1027
 Db 856 CGCAGGAAGATTCATATTCATATGAGAGCTTAATTAATACCGGCTTCTGAAGAT 915
 QY 1028 TCATTGGAGAAATGCTAAACCGTTATATTTGCAAAATTAAGCCAGTTTC-----T 1081
 Db 916 TCCTGGAGAGCATGCTAAGATGTCATGATCATAGTGTGACAGCCCTCCTCGAAT 975
 QY 1082 TTTGATGAGACTCTAATACACTTTCAGTTGCCAGTACTGCCAAACATGTGAGAAATACT 1141
 Db 976 TTTGATGAGCTTAAATTTCTCTCAAAATATGCCAAGAGACGGAACATTAGAAACAA 1035
 QY 1142 CCCCATGTAA 1152
 Db 1036 CCCACTGTAA 1046
 RESULT 2
 US-09-641-807A-1
 ; Sequence 1, Application us/09641807A
 ; Patent No. 6440731
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Freedman, Richard
 ; TITLE OF INVENTION: No. 6440731el motor proteins and methods for
 ; FILE REFERENCE: 1031
 ; CURRENT APPLICATION NUMBER: us/09/641,807A
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3837
 ; TYPE: DNA
 ; ORGANISM: Human
 ; NAME/KEY: misc_feature
 ; LOCATION: (1226)....(1336)
 ; OTHER INFORMATION: n = a, t, c, or g
 US-09-641-807A-1
 Query Match 1.5%; Score 147.4; DB 4; Length 3837;
 Best Local Similarity 51.3%; Pred. No. 1,6e-26;
 Matches 467; Conservative 0; Mismatches 411; Indels 33; Gaps 4;

QY 275 TTCAATTCGATCGTATTTAATTTCTCAGCAATCAACAGTCAATTTTACCAAGAAATA 334
 Db 136 TTCACTTTGATTTTGTGTTTGGCAAAATTCACCTCAAGATGAAGTTTAAACATGT 195
 QY 335 GCACTACCTATCATTACATCAGCTTTGCAAGGATATATGACCAATTTTGCATACGA 394
 Db 196 ATAAAGCCCCAGTGTGTCTCCTCATGAGGCTATATGCAACTGTTTGTGCTATGGA 255
 QY 395 CAGACATCTTCAGCAAGAGCTACACATATGATGGGAACCAATTCAT----- 444
 Db 256 CAACCTGATCTGGGAAGACATACACATTTGAGAGGGCCCATATTCCTTCAAGTTGTGAG 315
 QY 445 -----GGGCATATATACCCCAAGCCATACAGAGATTTTAAATTTATCAGAGATA 496
 Db 316 GGCCAAAGGATATCATTTCCAGATTTCAAGAAATTTCAAGCATCTGTGAACAT 375
 QY 497 CCGAAGAGAGTTTCTTCTTAAGGTTTCTTATTTGAGATTTCAATGAACTGTGAAA 556
 Db 376 CCTAGCATTTGACTTTAATGTAAGGATCTTATATGAAAGTATCAAGAGAACCTAAGA 435
 QY 557 GACCTACTGTGTATGACAGAGAAAGAGCCCTGGAAATTCGCGAGATTTTAATAGA 616
 Db 436 GATCTTTAGATTTGGAGACATCCATGAAGATTTTCAATCCGAGAGATGAAAGAA 495
 QY 617 AACGTGATGTGTGACTGACTGAGAAACTTGAATGTTTCTGTAACATGTATACAG 676
 Db 496 AACCACTGATATTTGGGGCCCAAGGAATGCCATGTGAGAGTGAAGTGAATGATAGT 555
 QY 677 TGGATCAAAAAAGGTGAAAAAAGACATTTATGAGAGACTAAATGATCATAGT 736
 Db 556 CTTTGGAGATGGGAAATGACCGACATACAGGTCCACTCAAAATATGAGACACTCC 615
 QY 737 AGTCGTTCACTACATATTTATTAAGATTTGTAAGCCGAGACAGAAAT-----GAT 790
 Db 616 AGCAGATCAGATGAATTTTACATAGACATTTGTCAATGCTAATTAATATGAGAGCA 675
 QY 791 CCCCAAAATTCAGAAAGTGTGATGAGCTGTCTATGCTATCTCAATTTGATAGT 850
 Db 676 GCTGAAGATGATCATGATGATTTCCCTCGGCATATTTCTCAAAAGTTCCACTTTGTGAT 735
 QY 851 CTTTGTGCAAGTGAAGAGCAAGCCAACTGAGCTGAGAGTGTGAGACTTAAGAAAGGC 910
 Db 736 TTGGCAGAGATCAGAAAGATTAACCAAAACGGGAATACTGTGACGTTCCAAAGAAATCC 795
 QY 911 TGCACATCAACCGCAGCTTGTATTCCTTGACAGGTTATTAAGAAAGCTTAGAGAGC-- 968
 Db 796 ATTCAAATCAATAGTATGATGCTGCTTTAGAGAAATTAATGAAGCGCTTGGAGACCA 855
 QY 969 -GCCAGCTGTGATTTATTAACACTACAGAGACGCAAACTCAGCAATTTCTCCAAAT 1027
 Db 856 CGCAGGAAGATTCATATTCATATGAGAGCTTAATTAATACCGGCTTCTGAAGAT 915
 QY 1028 TCATTGGAGAAATGCTAAACCGTTATATTTGCAAAATTAAGCCAGTTTC-----T 1081
 Db 916 TCCTGGAGAGCATGCTAAGATGTCATGATCATAGTGTGACAGCCCTCCTCGAAT 975
 QY 1082 TTTGATGAGACTCTAATACACTTTCAGTTGCCAGTACTGCCAAACATGTGAGAAATACT 1141
 Db 976 TTTGATGAGCTTAAATTTCTCTCAAAATATGCCAAGAGACGGAACATTAGAAACAA 1035
 QY 1142 CCCCATGTAA 1152
 Db 1036 CCCACTGTAA 1046
 RESULT 3
 US-09-723-096-1
 ; Sequence 1, Application us/09723096
 ; Patent No. 6448026
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Freedman, Richard
 ; TITLE OF INVENTION: No. 6448026el motor proteins and methods for


```

; TITLE OF INVENTION: their use
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/723,096
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3837
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1236)..(1337)
; OTHER INFORMATION: n = a, t, c, or g
US-09-723-096-1

```

```

Query Match      1.5%; Score 147.4; DB 4; Length 3837;
Best Local Similarity 51.3%; Pred. No. 1.6e-26;
Matches 467; Conservative 0; Mismatches 411; Indels 33; Gaps 4;

```

```

QY 275 TTCAATTGCGTCGTGTTAATTTCTACGATCAACAGTCAATTTACCAAGAAATA 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 TTCACTTTGATTTGTTGTTGGCAAAATTCCTCAAGATGAAGTTATACACATGT 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 335 GCAGTACCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 ATAAAGCCCCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 395 CAGACATCTTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 CAATGATCTGGGAAGACATACACATGAGGGGGCCATATGCTTCACTGTTGGAG 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 -----GGGCAATATACCCCAAGCCATACAGAGATTTTAAATTTATTCAGAGATA 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 GGGCAAAAGGATATCTCTGAGATTTCTATATGAGATTTTACATGAAGATGAGAA 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 497 CCGAAGACAGAGTTCTTCTAAGATTTCTTATATGAGATTTTACATGAAGATGAGAA 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 CTTGATCTGATCTTAAATGTAAGATATCTTATGAGATTTACAGAGATGAGAA 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 557 GACCTACTGTGTATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 GATCTTCTAGATTTGAGACATCCATGAGATCTTCACTCCAGAGATGAGAAAGAA 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 617 AACGTGTATGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 AACACAGTATGTTGGGCGCAAGAGATGCCATGTGAGAGTGCAGTGAAGTATGAGT 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 677 TGGATCAAAAAGGGTGAAGAAAGACATATGAGAGAGAGAGAGAGAGAGAGAGAG 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 556 CTTTGGAGATGGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 737 AGTGTTCATATACATATTTAGATGATTTGTAAGCCGAGAGAGAGAGAGAGAGAG 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 616 AGCAGATCAGATGAGATTTTACATGAGATTTGTAAGCCGAGAGAGAGAGAGAGAG 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 791 CCCCAAAATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 676 GCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 851 CTTGCTGCAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 736 TTGCGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 911 TGCACATCAACCGCAGCTGTTTATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 796 ATTCAATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 969 -GCCAGGCTGTGATTTATTAATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 856 CCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 1028 TCATTGGAGAGAAATGCTAAACGGTTATTAATTTGCAATATACGCACTTCT-----T 1081
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 916 TCTCTGGAGGAGAGCTTCTAAGACTGTGATGATGATGATGATGATGATGATGATGAT 975
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1082 TTTGATGAGACTCTAAGTACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 976 TTTGATGAGCTCTTAATTTCTCTCAAAATATGCAACAGAGAGAGAGAGAGAGAGAGAG 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1142 CCCCATGTTAA 1152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1036 CCCCCTGTTAA 1046
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
US-09-724-517-3
; Sequence 3, Application US/09724517
; Patent No. 6379941
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6379941el motor proteins and methods for
; FILE REFERENCE: 1031
; CURRENT APPLICATION NUMBER: US/09/724,517
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US/09/641,807
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 3
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Human
US-09-724-517-3

```

```

Query Match      1.5%; Score 140; DB 4; Length 1023;
Best Local Similarity 51.0%; Pred. No. 5.3e-25;
Matches 456; Conservative 0; Mismatches 405; Indels 33; Gaps 4;

```

```

QY 275 TTCAATTGCGTCGTGTTAATTTCTACGATCAACAGTCAATTTACCAAGAAATA 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 TTCACTTTGATTTGTTTGGCAAAATTCCTCAAGATGAAGTTATTAACATGT 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 335 GCAGTACCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 ATAAAGCCCTGATGTTGCTACTCATGAGGCTTAAATGCAACGTGTTTCTATGGA 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 395 CAGACATCTTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 CAATGATCTGGGAGAGACATACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 -----GGGCAATATACCCCAAGCCATACAGAGAGTTTAAATTTATTCAGAGATA 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 GGGCAAAAGGATATCTCTGAGATTTCTGAGATTTTCAAGAAATTTTCAAGCATCTCTG 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 497 CCGAAGAGAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAGAT 556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 CTTACATTTGATCTTAAATGTAAGATTTCTTAAAGATTTCTTAAAGATTTCTTAAAG 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 557 GACCTACTGTGTATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 GATCTTTCTAGATTTGAGAGAGATCTTCAAGAGATTTCTTCAATCTCGAGAGAGATGA 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 617 AACGTGTATGTTGCTGACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 490 AACACAGTATGTTGGGCGCAAGAGATCCATGTGAGAGAGAGAGAGAGAGAGAGAGAG 549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 677 TGGATCAAAAAGGGTGAAGAAAGACATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 CTTTGGAGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 737 AGTGTTCATATACATATTTAGATGATTTGTAAGCCGAGAGAGAGAGAGAGAGAGAG 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 610 AGCAGATCAGATGATTTTACATATGATGATTTGTCAGTTGCTAATAAATATGAGAGCA 669
QY 791 CCCACAAATTCAGAGAACTGTGATGAGCTGTCTATCTCATCTTGAATTTGGTAGAT 850
Db 670 GCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
QY 851 CTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910
Db 730 TTGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
QY 911 TGCAACATCAACCGCAGCTGTTTATCTTGAAGAGTTTAAAGAGTTTAAAGAGTTTAA 968
Db 790 ATTCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
QY 969 -GCCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
Db 850 CGCAGGAGAGTTTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 909
QY 1028 TCATTTGGAGGAAATGCTTAAACGTTATATTTGACAAATTTACCGAGTTTC-----T 1081
Db 910 TCTCTGGAGAGCAGTCTTAAAGCTGTATGATGATGATGATGATGATGATGATGATGAT 969
QY 1082 TTTGATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
Db 970 TTTGATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023

RESULT 5
US-09-641-807A-3
Sequence 3, Application US/09641807A
Patent No. 6440731

GENERAL INFORMATION:
APPLICANT: Berard, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6440731el motor proteins and methods for
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/641,807A
CURRENT FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1023
TYPE: DNA
ORGANISM: Human
US-09-641-807A-3

Query Match 1.58; Score 140; DB 4; Length 1023;
Best Local Similarity 51.08; Pred. No. 5.3e-25;
Matches 456; Conservative 0; Mismatches 405; Indels 33; Gaps 4;

QY 275 TTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 334
Db 130 TTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 189
QY 335 GCAGTCTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 394
Db 190 ATAAAGCCCTAGTGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 249
QY 395 CAGACATCTTACAGCAAGCTTACAGCAAGCTTACAGCAAGCTTACAGCAAGCTTACAGCA 444
Db 250 CAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 309
QY 445 -----GGCATAATATCCCAAGCCATACAGCAAGTTTAAATATTTACAGAGATA 496
Db 310 GGCATAATATCCCAAGCCATACAGCAAGTTTAAATATTTACAGAGATA 369
QY 497 CGGACAGAGAGTTTCTTAAAGTTTCTTAAAGTTTCTTAAAGTTTCTTAAAGTTTCTTAAAG 556
Db 370 CCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429
QY 557 GACCTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 616

Db 430 GATCTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
QY 617 AACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
Db 490 AACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
QY 677 TGATCAAAAAGGTTGAAAAAGGTTGAAAAAGGTTGAAAAAGGTTGAAAAAGGTTGAAAA 736
Db 550 CTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
QY 737 AGTCTCTACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
Db 610 AGCAGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
QY 791 CCCACAAATTCAGAGAACTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 850
Db 670 GCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
QY 851 CTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910
Db 730 TTGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
QY 911 TGCAACATCAACCGCAGCTGTTTATCTTGAAGAGTTTAAAGAGTTTAAAGAGTTTAA 968
Db 790 ATTCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
QY 969 -GCCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
Db 850 CGCAGGAGAGTTTCATATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 909
QY 1028 TTTGATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
Db 910 TCTCTGGAGAGCAGTCTTAAAGCTGTATGATGATGATGATGATGATGATGATGATGAT 969
QY 1082 TTTGATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1135
Db 970 TTTGATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023

RESULT 6
US-09-723-096-3
Sequence 3, Application US/09723096
Patent No. 6448026

GENERAL INFORMATION:
APPLICANT: Berard, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6448026el motor proteins and methods for
FILE REFERENCE: 1031
CURRENT APPLICATION NUMBER: US/09/723,096
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: US/09/641,807
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1023
TYPE: DNA
ORGANISM: Human
US-09-723-096-3

Query Match 1.58; Score 140; DB 4; Length 1023;
Best Local Similarity 51.08; Pred. No. 5.3e-25;
Matches 456; Conservative 0; Mismatches 405; Indels 33; Gaps 4;

QY 275 TTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 334
Db 130 TTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 189
QY 335 GCAGTCTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 394
Db 190 ATAAAGCCCTAGTGTGTGTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 249

```

OY 395 CAGACATCTTCAGCAGACGATACACATGATGGAAACCAATTCATT----- 444
DB 250 CAAATCGATCTGGGAAAGCATACACATGGAGGGGCCATATTCCTTCAGTTGGAG 309
OY 445 -----GGGCAATATACCCCAAGCCATACAGAAATTTTAAATATTCAGAGATA 496
DB 310 GGCACAAAGGATATCTTCCTCGAGCTATTCAGAAATATTCAGAAATCTCTGAAAT 369
OY 497 CGGACAGAGAGTTCTCTTAGAGATTCTTATATGAGATTATACATTAACATGGA 556
DB 370 CCTACATGACTTAAATGTAAGATCTTATATGAGAGTACAGAAACCACTTAA 429
OY 557 GACCTACTGTGTGATGACAGAAAGAGCCCTTGAATTCGGAGATTTTATAGA 616
DB 430 GATCTCTAGAAATTTGAGACATCCATGAAGATCTTCACATCCGAGAGATGAAAGGA 489
OY 617 AACGTGTATGTTGCTGACCTGACAGAACTTGTATGTTCTGTAACTATATACG 676
DB 490 AACACAGTATGTTGGGGCCAAAGAAATCCATGTGAGAGTGCAGGTGAAGTATGAGT 549
OY 677 TGGATCAAAAAGGTGAAAAAACAGACATTTATGAGAGACATTAATGATCATAGT 736
DB 550 CTTTGGAGATGGGAAATGACAGCAGACATACAGGTACCACTCAATGATGAGACTCC 609
OY 737 AGTCCTGACATACATATATTTAGATGATTTGAAAGCCGAGACAGAAAT-----GAT 790
DB 610 AGCAATATCATGCAATTTTACATCAGCATTTGTCAAGTTCATTAATAATGAGAGCA 669
OY 791 CCCACAAATTCAGAGAACTGTGAGAGCTGTATGATCTGATCTGAAATTTGGTAGAT 850
DB 670 GCTGAGATGATGATGATGATTCCTCGGCAATATTTGTCATAAGTTCACACTTGTGAT 729
OY 851 CTGTGTCAGAGAAAGCAAGCAAACTGAGCTGAAGTGTGAGACTTAAAGAGAGC 910
DB 730 TTGGAGAGATCAGAAAGTAACCAAAACGGGAAATCTGAGGAAAGGTTTCAAGAAATCC 789
OY 911 TGCAACATCAACCCGAGCTGTTTATCTTGGACAGGTATTAAGAAGCTTATGACAGC-- 968
DB 790 ATTCAATCAATAGTGAATGCTGCTTAGGAATGTATTAAGGCTCTTGGGACCA 849
OY 969 -GCCAGCTGTGTATTAATAACTACAGAGACAGCAAACTGACAGAAATTCCTCAAAAT 1027
DB 850 CGCAGAGAGATGATCAATTCATATAGGATGCTAAATTAACCCGGCTCTGAAAGAT 909
OY 1028 TCATGAGAGAAATGCTAAACGTTTAAATTTGCACATTAACCCAGTTT-----T 1081
DB 910 TCTCTGGAGCAGTGTAGACTGTCAATGATCAGATGTGAGCCCTCCCTCTCGAAT 969
OY 1082 TTTGATGAGACTCTAAGTACACTTCACTTGGCAGTACTGCCAAACATGTGAGA 1135
DB 970 TTTGATGAGCTCTTAATTTCTCTCAAAATATGCAACAGACAGCAACATTTAGA 1023

```

```

RESULT 7
US-09-592-054-5
: Sequence 5, Application us/09592054
: Patent No. 6440684
: GENERAL INFORMATION:
: APPLICANT: Berard, Christophe
: APPLICANT: Finer, Jeffrey
: APPLICANT: Sakowicz, Roman
: APPLICANT: Wood, Kenneth
: TITLE OF INVENTION: No. 6440684el motor proteins and methods for
: TITLE OF INVENTION: their use
: FILE REFERENCE: 1016
: CURRENT APPLICATION NUMBER: us/09/592, 054
: CURRENT FILING DATE: 2000-07-20
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 1421
: TYPE: DNA

```

```

: ORGANISM: Human
US-09-592-054-5

```

```

Query Match 1.4%; Score 135.2; DB 4; Length 1421;
Best Local Similarity 50.6%; Pred. No. 9,8e-24;
Matches 468; Conservative 0; Mismatches 408; Indels 48; Gaps 4;

```

```

OY 257 GTTGATGGACAAAGCTTTCATTTGATCGTGTATTAATTCACAGATCAACAACT 316
DB 123 GTTGGTACAGATTAATCTTTCACCTACGATTTTGTATTTGATCCCTCTCTGAAACGGA 182
OY 317 CAAATTTACCAAGAAATACAGTACCTATACATACATCAGTTTGGAGGATATATGCG 376
DB 183 GAAGTCTCAATACAGACATAGCGGCACATCAATAAAGGTATTTAAAGATATATATGA 242
OY 377 ACAATTTTTCATAGGAGACATCTTACGACAGACGTACACATATGATGGAAACCA 436
DB 243 ACGTCTCGGCTATGGGACACTGGCTCTGAAAAAACCATTATTCATGAGAGTGCATAT 302
OY 437 AATTCATG-----GGCATATACCCCAAGCCATACAGAAAT 475
DB 303 ACTGACAGCAAGAGATGACCAACAGTTGGGTTATTTCTTAGGTTATACACTGCTC 362
OY 476 TTTAAATTTATTCAGAGATACCGAACAGAGATTTCTTAAGATTTCTTAATGAG 535
DB 363 TTCAAAAGAAATGATTAAGAGAGTACTTTGAATTTACTGAAAGTGTCTTAAGTAG 422
OY 536 ATTTACATGAACATGTGAAGAACTACTGTGTATGACAGAAAGAAAGCCCTTGGCA 595
DB 423 ATTTCAAAAGAAATTTTGAATCTTATCCATCCATCTCGAGAAACCTCAATTAAT 482
OY 596 ATTCGAGAGATTTTAAATGAACGTAATGTTGCTGACCTGACATGAACCTGTGAATG 655
DB 483 ATACGAGAGATCTTAAGAGAGCATTAAGATTTGAGACTACTGAGAAAGCTGTTTGG 542
OY 656 GTTCTGAAATGTAATACATGATGATCAAAAAGGTTGAAAAAAGACATTAATGAGAG 715
DB 543 GTTGCCTTGATACGTGTTCTGTTTGGAAACGGGCAACACTAGAGACTGTGGGCTCC 602
OY 716 ACTAAATATGATCATATGATGCTGTTGACATACATATTTAGAAATGATTTGAAAGC 775
DB 603 ACGGTATGTAACCTCCAGTCTCCGATCTCATGTCATCTTTTCAAT--CTCTTACAGC 660
OY 776 CGAGACAGAAATGATCCCAATTCAGAACTGATGAGAGCTGTATGATCTGC 835
DB 661 AAAGAAAGAAAGTGCAGAAATAGCAG-----CTTTGCTTCAAG 701
OY 836 TTGAATTTGATGATCTTGTGCTGAGTGAAGAGCAAGCCAACTGAGCTGAGGTGTG 895
DB 702 CTGCAATCTTGTAGACTGCTGCTGATCAGAAAGACAGAAAGAAACCAAGGCTGAAGGGAT 761
OY 896 AGACTTAAGAAAGCTGCAACATCAACCGCAGCTGTTTATCTTGGACAGGTTTAAAG 955
DB 762 CECTTAAGAGAGGTATTAATTAATTAACCGAGCTCTATGCTTGGAAATGTAATCAAT 821
OY 956 AAGCTTAGAGCAGCGCAGCTGTGATTTATTAATCTACAGACAGCAAACTGCACGA 1015
DB 822 GCTCTTGGAGATGACAAAAAGGCTGTGTTGCTTACAGAGATTTCCAAAGTTGACTGA 881
OY 1016 ATTTCCAAATTCATTTGGAGAAATGCTAAAGGTTATTAATTTGCAATTAAGCCA 1075
DB 882 CTGCTTCAAGATTTCTAGAGAGTAAATAGCATATCTTATGATAGCGTGTGAGTCT 941
OY 1076 GTT-----TCTTTGATGAGACTTAAGTACATCTTCACTTGGCAGTACTGCCAAACT 1129
DB 942 GCTGACTCCAACTAGAGAAACATTAATACCTTGTCTAGCTGACAGACAGCAAGAAA 1001
OY 1130 GTGAGAAATATCTCCCATGTTAAT 1153
DB 1002 ATCAAGAACAACTATTTGTTAAT 1025

```

```

RESULT 8

```

US-09-592-054-3
 ; Sequence 3, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Beraud, Jeffrey
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684e1 motor proteins and methods for
 ; FILE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592,054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1569
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-592-054-3

Query Match 1.48; Score 135.2; DB 4; Length 1569;
 Best Local Similarity 50.68; Pred. No. 1e-23;
 Matches 468; Conservative 0; Mismatches 408; Indels 48; Gaps 4;

QY 257 GTTGATGGACAAAGCTCTTCAATTCGATCGTATTTAATTCACGAATCAACAAGT 316
 DB 187 GTTGATGAGATTAATCTTCAACCTACCATTTTGTATTTGATCCCTACGAAACAGAA 246
 QY 317 CAAATTTACCAAGAAATAGCAGTACCTATCATAGCATGACTTTGACAGGATATATGCG 376
 DB 247 GAAGCTTCAATACAGCAGTACGCGCTCATAAAGGTGATTAAGGATATATATCA 306
 QY 377 ACAATTTTGCATAGGAGACATCTTCAGCAAGAGCTACACATATGATGGAACACCA 436
 DB 307 ACGGTCTGCGCTATGGCAGACTGCTGTGAAAAACCTATCATATGAGGATGACATAT 366
 QY 437 AATTCATTG-----GGCATATATACCCCAAGCCATACAGAACTT 475
 DB 367 ACTGACAGACAGAGAAATGAAACCAACAGTGGGGTATTCCTAGAGGATATCAACTGCTC 426
 QY 476 TTTAAATTTATTCAGAGATACCGAAGAGAGTTCTTCTAAGAGTTCTTATATGAG 535
 DB 427 TTCAAGAAATTTGATTAAGAGAGTGTGATTTACTGTAAAGTGTCTTACTTAAGAG 486
 QY 536 ATTACAAATGAAGTGTGAAGAGCTACTGTGTGATGACAGAAAGAAAGAGCCCTTGGAA 595
 DB 487 ATTACAAATGAAGAAATTTTGGATCTTGTATGCCATCTCGAAGAAAGCTCAATTAAT 546
 QY 596 ATTCGCGAGATTTTATAGAAAGCTGTATGTTGCTGACCTGACTGAAAGAACTTATATG 655
 DB 547 ATACGAGAGATCTTAAGAGAGCATTAAGATTGTGGGCTCAGCAGAGAAAGACTTTTGG 606
 QY 656 GTTCCGAGACATGTATACAGTGTGATGATCAAAAGGGTGAAGAAACAGACATATGAGAG 715
 DB 607 GTTGCTTGGATCTGTCTTCTGTGTGAGACAGGGCAACAACCTTAGAGCTGTGCTTCC 666
 QY 716 ACTAAATATGATGATAGTAGTGTCTTCAATATATTAATTTAGATGTTTGAAGAG 775
 DB 667 ACGGCTATGATCCAGTCGTCGCGATCTCATAGCATTTTACAT--CTCCTTAGAGC 724
 QY 776 CGAGACAGAAATGATCCCAAAATTCAGAGAACTGTATGAGCTGTATGTATCTAC 835
 DB 725 AAAGAAAGAAAGTACAGAAATAGAG-----CTTTCGCTCCAG 765
 QY 836 TTGAATTTGATCTGTGCGAGTGAAGAGCAAGCCAAACTGAGCTGAGAGTGTG 895
 DB 766 CTGACATCTGTGACCTCTGCTGATGAGAAAGCAGAAACCAAGGCTTAAGAGGAT 825
 QY 896 AGACTTAAGAGAGCTGCAACATCAACCCAGCTGTTTATCTTATCTTGACAGGTTTAAAG 955
 DB 826 CGCTTAAGAGAGGATTTATATTAACGAGAGGCTCTTGTGTGGAAATGATTAACAT 885

QY 956 AAGCTTAGCAGGCGCCAGCTGTGATTTATTAATCTACAGAGACGAACCACTACCA 1015
 DB 886 GCCTCTGAGATGACAAAGAGGTGGCTTTGTGCTTACAGAGATTCAGTGA 945
 QY 1016 ATTCCCAAAATTCATTGGAGAGAAATGCTAAACGGTATATATTTGCACATTTGCGCA 1075
 DB 946 CTGCTTCAAGATCTCTAGAGAGTAAAGCCATACCTTATATATAGCTGTGTGAGTCT 1005
 QY 1076 GTT-----TCTTTGATGAGACTATAGTACACTTCACTTGTGCGAGTACTGCCAAACAT 1129
 DB 1006 GCTGATCCCAATCTAGAGAGAAATTAATTAACCTTCCGTATGCTGAGACAGCAAGAAA 1065
 QY 1130 GTGAGAAATACCTCCCATTTAT 1153
 DB 1066 ATCAAGAACCAACCTATTTAT 1089

RESULT 9
 US-09-592-054-7
 ; Sequence 7, Application US/09592054
 ; Patent No. 6440684
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Beraud, Jeffrey
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6440684e1 motor proteins and methods for
 ; FILE OF INVENTION: their use
 ; FILE REFERENCE: 1016
 ; CURRENT APPLICATION NUMBER: US/09/592,054
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 4127
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-592-054-7

Query Match 1.38; Score 127.2; DB 4; Length 4127;
 Best Local Similarity 50.18; Pred. No. 1.6e-21;
 Matches 463; Conservative 0; Mismatches 413; Indels 48; Gaps 4;

QY 257 GTTGATGGACAAAGCTCTTCAATTCGATCGTATTTAATTCACGAATCAACAAGT 316
 DB 229 GTTGATCTGATTAATCTTCAACCTACCATTTTGTATTTGATCCCTACTGAGCAGAA 288
 QY 317 CAAATTTACCAAGAAATAGCAGTACCTATCATAGCATGACTTTCAGAGGATATATGCG 376
 DB 289 GAATCTTCAATTAAGAGAGTACGCGCTCATAAAGCATATTTAAAGATTAATATGCA 348
 QY 377 ACAATTTTGCATAGGAGACATCTTCAGGCGCAAGCGTACACAAATGATGGAACACCA 436
 DB 349 ACGGTCTGCGCTATGGCAGACTGCTGTGAAAAACCTATTCATGAGAGGTGACATAC 408
 QY 437 AAT-----CATTGGCATATATACCCCAAGCCATACAGAAAGT 475
 DB 409 ACTGCGAGACAGAGATGAAACCAACAGTGGCATTTTCTAGGATATACCAATGCTC 468
 QY 476 TTTAAATTTATTCAGAGATACCGAAGAGAGTTCTTAAAGATTTTATATAGAG 535
 DB 469 TTCAAGAAATTTGATCAAAAGAGTGTGATTTACTGTGAAGAGTGTCTTACTTAAGAG 538
 QY 536 ATTACAAATGAAGTGTGAAGAGCTTGTGTGATGACAGAAAGAAAGAGCCCTTGGAA 595
 DB 529 ATTACAAATGAAGAAATTTTGGATCTTATGCCATCTGTGAGAAAGCTCAATTAAT 588
 QY 596 ATTCGCGAGATTTTATAGAAAGCTGATGTTCTGACCTGACTGAGAAAGCTTATATG 655
 DB 589 ATACGGGAGATCTTAAGAGAGCATTAAGATTTGTGGACCTCACTGAGAGAGCTTTTA 648
 QY 656 GTTCCGAGACATGTATACAGTGTGATCAAAAGGTTGAAAAACAGACATATATGAGAG 715

```

Db 649 GTTGCCCTGGATCTGTTCTCTGTTGGACAGGCGACAACTCTAGAGCTGCGCTCC 708
QY 716 ACTAAATGATGATCATCTAGTACGTTTACATACATATTTAGATGATGTTGAAGC 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 ACAGCATGAGTACCCAGGTGTCGCCATCTCATGCCATCTTTACAA--CTCCTTAGAGC 766
QY 776 CGAGACGAATATATCCCAATTCAGAGACTGTGATGAGCTGTCAATGATCTCAC 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 767 AAGGAAGAAAGTACAAAGATAGCAG-----CTTTCGCTCCAG 807
QY 836 TTGAATTTGGATATCTGCTGGCAGTGAAGAGCAAGCCAACTGAGTGAAGGTG 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 808 CTCATCTTGTAGACCTCGGTGATCAGAAAGACAGAAAGCAAGCTGAAGGGGAT 867
QY 896 AGACTTAAGAGAGCTGCAATCAACCCGAGCTGTTTATCTCTGGACAGTTATTAAG 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 868 CGCTAAAGAGGGTATTAATATTAACCGAGGCTCCTATGCTTGGAAATGTAATAGT 927
QY 956 AAGCTTAGCGAGCGCGAGCTGTGATTTATAACTACAGAGACAGCAAACTCACCA 1015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 928 GCTCTTGAGANTGACAAAAGGGTAGCTTGTGCCCTACAGAGATTCCAAGTTAACTCGA 987
QY 1016 ATTCTCCAAATTCATTTGGAGGAATGCTAAACGGTTATTTTGCACAAATTAAGCCA 1075
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 988 CTGCTGAGATCTCTAGAGAGGTACAGCCACACTCTTATGATAGCTGTGAGTCT 1047
QY 1076 GTT-----TCTTTGATGAGACTTAAGTACACTTCAGTTGCTCCAGTACTGCCAAAT 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1048 GCGACTCCCAATCTAGAGGAACATTAAGTACCTTCGCTATGCTGACAGAGCAAGAAA 1107
QY 1130 GTGAGAAATACTCCCATGTTAAT 1153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1108 ATCAAGAACAAACCTATTTGTTAAT 1131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
US-09-592-054-1
; Sequence 1, Application US/09592054
; Patent No. 6440684
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Finer, Jeffrey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6440684e1 motor proteins and methods for
; FILE REFERENCE: 1016
; CURRENT APPLICATION NUMBER: US/09/592,054
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4308
; TYPE: DNA
; ORGANISM: Human
US-09-592-054-1

```

```

Query Match 1.2% Score 116; DB 4; Length 4308;
Best Local Similarity 49.4%; Pred. No. 1e-18;
Matches 456; Conservative 0; Mismatches 420; Indels 48; Gaps 4;

```

```

QY 257 GTTGATGGAGCAAGTCTTCAATTCGATCGTATTTAATTCGACGAATCAACAAGT 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 GTTGTGATAGATTAATCTTACCTACGATTTTGTATTTGATCTCCCTACTGAACAGAA 246
QY 317 CAATTTACCAAGAAATAGCAGTACATATCATACGATCAGCTTTGCAAGGATTAATGCC 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 GAAGTCTTAATATACAGCAGTAAAGCCATCAATAAAGGTGATTAAGATTAATATGCA 306
QY 377 ACAATATTTGATAGCAGACATCTTCAGCAAGACGATCAATGATGGAACCA 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 ACGGTCTGCGCTATGAGGAGACTGGCTGGAAGAAACCTATTTCAATGGAGGTGATAT 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 437 AATTCATG-----GGCATTAATACCCCAAGCCATACAGAGATT 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 ACTGACAGCAAGAGAAATGACCAACAGTTGGGGTTATTCCTAGGGTAATACAACTGCTC 426
QY 476 TTTAAATTTATTCAGGATACCGAACAGAGATTCTTTAGAGTTCTTTATAGAG 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 TTCAAAGAAATGTTAAAGAGTGAATTTGATTTTACTGTGAAGTGTCTTACTTAGAG 486
QY 536 ATTTACATGAACTGTGAAGACCTTACTGTGATGACAGAAAGAAAGACCTTTGAA 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 ATTTACATGAAAGAAATTTTGATCTTATGTCATCTGTGAGAAAGCTCAATTAGAG 546
QY 596 ATTCGCGAGATTTTAAATAGAAACGTGATGTTGCTGACCTGACTGAAGACTTTAATG 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 ATTACAAATGAAAGAAATTTTGATCTTATGTCATCTGTGAGAAAGCTCAAGTTTG 606
QY 656 GTTCCCTGAAATGTAATACAGTGAATCAAAAGGGTGAAGAAAGACATTTATGAGAG 715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 GTTGCCCTTGATCTGTTCTCTGTTGGAAACAGGCAACCTGTAGACTGTGCGCTCC 666
QY 716 ACTAAATGATGATCATCTAGTACGTTTACATACATATTTAGATGTTTGAAGC 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 ACGGCTATGAACTCCAGTGTGCCATCTCATGCCATCTTAACAA--CTCCTTAGAGC 724
QY 776 CGAGACGAATGATCCCAAAATTCAGAGACTGTGATGAGCTGTATCTCAC 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 AAGGAAGAAAGTACAAAGATAGCAG-----CTTTCGCTCCAG 765
QY 836 TTGAATTTGGTATATCTGCTGGCAGTGAAGAGCAAGCCAACTGAGTGAAGGTG 895
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 CTGATCTTTAGACCTCGGTGATCAGAAAGACAGAAAGCAAGCTGAAGGGGAT 825
QY 896 AGACTTAAGAGAGCTGCAATCAACCCGAGCTGTTTATCTCTGGACAGTTATTAAG 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 CGCTAAAGAGGGTATTAATATTAACCGAGGCTCCTATGCTTGGAAATGTAATCAGT 885
QY 956 AAGCTTAGCGAGCGCGAGCTGTGATTTATAACTACAGAGACAGCAAACTCACCA 1015
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 GCTCTTGAGANTGACAAAAGGGTAGCTTGTGCCCTACAGAGATTCCAAGTTGACTCGA 945
QY 1016 ATTCTCCAAATTCATTTGGAGGAATGCTAAACGGTTATTTTGCACAAATTAAGCCA 1075
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 946 CTGCTTCAAGATCTCTAGAGAGGTAAAGCCATACCTTATGATAGCTGTGAGTCT 1005
QY 1076 GTT-----TCTTTGATGAGACTTAAGTACACTTCAGTTGCTCCAGTACTGCCAAAT 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1006 GCGACTCCCAATCTAGAGGAACATTAATACCTTCGCTATGCTGACAGAGCAAGAAA 1065
QY 1130 GTGAGAAATACTCCCATGTTAAT 1153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1066 ATCAAGAACAAACCTATTTGTTAAT 1089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 11
US-09-641-806-3
; Sequence 3, Application US/09641806
; Patent No. 6395527
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: No. 6395527e1 motor proteins and methods for
; FILE REFERENCE: 1034
; CURRENT APPLICATION NUMBER: US/09/641,806
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Human
US-09-641-806-3

```

Query Match 1.0%; Score 98; DB 4; Length 1011;
 Best Local Similarity 48.3%; Pred. No. 1.3e-14;
 Matches 346; Conservative 0; Mismatches 355; Indels 15; Gaps 2;

359 TTGCAGGATATTAATGACACATATTTGCATACGACAGACATCTTCAGGCAAGACGTAC 418
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 214 TTGCAGGCTTCATGCTCCACTGTCTTGTGCTATGTGTCACAGCGGGCTCAGGGAACATAC 273

QY 419 ACAATGATGGGAACACCA-----ATTCAITGGGCATTAATACCCAGCCATA 466
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 274 ACCATGAGGGAGGCCACCTCCCTCTTGAGATGAGCAGGCGATGTCTCCGAGGCCATG 333

QY 467 CAGGAATTTTAAATTTTACAGAGATACCGAACAGAGATTTCTTCTTAAGATTTCT 526
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 334 GCGAGGCGCTTCAGCTCATGATGAAACGACCTGCTTGACTGCTGTATCATGTGTCC 393

QY 527 TATATGAGATTTACATTAATGAACTGTGAAGACCTACTGTGTGATGACAGAAAGAG 586
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 394 TACCTGGAAGTGTACAGAGAGAGATTCGAGACCTGCTGAGGTGGGCACTGCCAGCCGT 453

QY 587 CCCTTGGAAATTCGCGAGGATTTTAATGAAGCGTGTATGTGTGCTGACCTGACAGAGAA 646
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 454 GACATCCAGCTCCGGAGATGAGCGCGGATGTGTGTGCTGTGCGGGGTGAAGAGGTC 513

QY 647 CTGTAAATGTTCTGGAACATGTAATACAGTGTACATCAAAAAGGTTGAAAAACAGACAT 706
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 514 GACGTGAGGCGCTGATGATGAGTGTGAGCTCTGAGATGGGCAACGCGCGCGGCAC 573

QY 707 TATGAGAGACTTAATGATGATCATAGTGTGCTGATACATATTAATTAAGATGAT 766
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 574 ACGGAGCGACGACCTCAACCACTGTCTAGCGGCTCACACAGGCTTTACCTGACCC 633

QY 767 GTTGAACCCGAGACAGAAATGATCCCAAAATTCAGAACTGTGATGAGCTGTCTATG 826
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 634 CTGGAGCAGCGGGGGCGGCCCAAGCCCTACCCCGCCCGCCGCGGACAGTGTCT 693

QY 827 GTATCTCACTGAATTTGTGATCTTCTGCACTGTAAGAGCAAGCAACCTGAGCT 886
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 694 GTCTCCAGTTCACCTGTGTGACTGCGGGCTCAGAGAGGGTCTTAACGCGGACG 753

QY 887 GAAGGTGAGACTTAAAGAGGCTGCACATCAACCGCAGCTGTTATCTTGGACAG 946
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 754 ACCGCGAGCGGCTCAAGAGAGATCCAGATCAACAGCAGCTCTGCGGCTGGCAGC 813

QY 947 GTTATTGAAGCTTAAGAGGCGGCGAGGCTGTGATTT---ATTAACATCAGAGACG 1003
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 814 GTCATCAGCGCCCTGGGGAGCCTCAGCGCGCGGCGGACCAATACCTACCGGACTCC 873

QY 1004 AAACCTACCAAGATTTCCAAATTCATTTGGAGAAATGCTAAACGTTATAT 1059
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 874 AAGATCACCGGATCTCAAAAGACTCGCTGGCGGGAACCGCAAGCGGTGATGAT 929

RESULT 12
 US-09-641-806-1
 ; Sequence 1, Application US/09641806
 ; Patent No. 6395527
 ; GENERAL INFORMATION:
 ; APPLICANT: Berand, Christophe
 ; APPLICANT: Freedman, Richard
 ; TITLE OF INVENTION: No. 6395527el motor proteins and methods for
 ; FILE REFERENCE: 1034
 ; CURRENT APPLICATION NUMBER: US/09/641,806
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1026
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-641-806-1

Query Match 1.0%; Score 98; DB 4; Length 1026;
 Best Local Similarity 48.3%; Pred. No. 1.3e-14;
 Matches 346; Conservative 0; Mismatches 355; Indels 15; Gaps 2;

359 TTGCAGGATATTAATGACACATATTTGCATACGACAGACATCTTCAGGCAAGACGTAC 418
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 217 TTGCAGGCTTCATGCTCCACTGTCTTGTGCTATGTGTCACAGCGGGCTCAGGGAACATAC 276

QY 419 ACAATGATGGGAACACCA-----ATTCAITGGGCATTAATACCCAGCCATA 466
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 277 ACCATGAGGGAGGCCACCTCCCTCTTGAGATGAGCAGGCGATGTCTCCGAGGCCATG 336

QY 467 CAGGAATTTTAAATTTTACAGAGATACCGAACAGAGATTTCTTCTTAAGATTTCT 526
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 337 GCCAGGCGCTTCAGCTCATGATGAGAGACCTGCTGTGCTGTGTGATCATGTGTCC 396

QY 527 TATATGAGATTTACATTAATGAACTGTGAAGACCTACTGTGTGATGACAGAAAGAG 586
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 397 TACCTGGAAGTGTACAAAGAGAGAGATTCGAGACCTGCTCGAGGTGGGCACTGCCAGCCGT 456

QY 587 CCCTTGGAAATTCGCGAGGATTTTAATGAAGCGTGTATGTGTGCTGACCTGACAGAA 646
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 457 GACATCCAGCTCCGGAGATGAGCGCGGATGTGTGTGCTGTGCGGGGTGAAGAGGTC 516

QY 647 CTGTAAATGTTCTCTGAACATGTAATACAGTGTACATCAAAAAGGTTGAAAAACAGACAT 706
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 517 GACGTGAGGCGCTGATGATGAGTGTGAGCTCTCTGAGATGGGCAACGCGCGCGGCAC 576

QY 707 TATGAGAGACTTAATGATGATCATAGTGTGCTGATACATACATTAATTAAGATGAT 766
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 577 ACGGAGCGACGACCTCAACCACTGTCTAGCGGCTCACACAGGCTTTACCGGAGAC 636

QY 767 GTTGAACCCGAGACAGAAATGATCCCAAAATTCAGAACTGTGATGAGCTGTCTATG 826
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 637 CTGGAGACGCGGGGGCGGCCCAAGCGCTACCCCGCCCGCCGCGGACAGTGTCT 696

QY 827 GTATCTCACTGAATTTGTGATCTTCTGCGAGTGAAGCAAGCAACCTGAGCT 886
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 697 GTCTCCAGTTCACCTGTGTGACTGCGGGCTCAGAGAGGGTGTCTCAAGAGGGCAGC 756

QY 887 GAAGGTGAGACTTAAAGAGGCTGCACATCAACCGCAGCTGTTATCTTGGACAG 946
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 757 ACCGCGAGCGGCTCAAGAGAGATCCAGATCAACAGCAGCTCTGCGGCTGGCAGC 816

QY 947 GTTATTGAAGCTTAAGAGGCGGCGAGGCTGTGATTT---ATTAACATCAGAGACG 1003
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 817 GTCATCAGCGCCCTGGGGAGCCTCAGCGCGCGGCGGACCAATACCTACCGGACTCC 876

QY 1004 AAACCTACCAAGATTTCCAAATTCATTTGGAGAAATGCTAAACGTTATAT 1059
 || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 877 AAGATCACCGGATCTCAAAAGACTCGCTGGCGGGAACCGCAAGCGGTGATGAT 932

RESULT 13
 US-09-724-519-5
 ; Sequence 5, Application US/09724519
 ; Patent No. 6414121
 ; GENERAL INFORMATION:
 ; APPLICANT: Wood, Kenneth
 ; APPLICANT: Finer, Jeffrey
 ; APPLICANT: Berand, Christophe
 ; APPLICANT: Mak, John
 ; TITLE OF INVENTION: Methods of screening for modulators of
 ; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
 ; FILE REFERENCE: 1014A
 ; CURRENT APPLICATION NUMBER: US/09/724,519
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 09/592,037
 ; PRIOR FILING DATE: 2000-06-12
 ; PRIOR APPLICATION NUMBER: 09/428,156
 ; PRIOR FILING DATE: 1999-10-27

```

? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FASTQ for Windows Version 4.0
? SEQ ID NO: 5
? LENGTH: 1542
? TYPE: DNA
? ORGANISM: Human
? OS: 09-724-519-5

```

Query Match	1.0%	Score 95.8	DB 4	Length 1542
Best Local Similarity	53.5%	Pred. No. 5.8e-14		
Matches 286; Conservative	0	Mismatches 222;	Indels 27;	Gaps 3.

OY	084	AAAAGGGTGAATAAAGGACATTAATGGAGAGCATTAAATGATGATCAATATATATCTTT	743
Dd	644	AAAAAGGGGGGCGAAAAAAGACAACACTGCAGCTACTGTGTGTAATCATACTTCATCTGT	703
OY	744	CACATCAATATTTAGATATCATATTTGTTGAAGCCGAGACAGAATATGCCAACAAATTAG	803
Dd	704	CCCACATCAGTTTTCTCTGTTACATCATATATGAMAA-----AACTACGA	748
OY	804	AGAACTGTATGAGGCTGTCATGATCTCACTGTAATTTGGTAGATCTTGTCGCAGT	863
Dd	749	TTCATGAGAGAAGAGCTTGTAAATTCGAAATGTGAATCTGTTGATCTTGAGAGAAAGT	808
OY	864	AAAGAGCAACCACAATCGAGCTGAGCGTAAGGTGTAGACTTAAGAAAGCTGCACAATCAC	923
Dd	809	AAAAATTTGGCCCTTGTGAGCTGTGTAATGAAGAGAGCTGGGAACTGGAATATTAATTC	868
OY	924	GCAAGCTTTTATCTTTTGACAGGTTATTAAGAACTATAGCACGCCAGCGCTGGTGAT	983
Dd	869	AATCCGCTTGACTTTGGGAAGGCTATTAACGCCCCCTGTATGAAGAACACCTCATG---	925
OY	984	TTATTAACATACAGACAGCAAACTCACCAGAAATTTCTCCAAATTCATTTGGAGGAATG	1043
Dd	926	--TTCCTTATTCAGAGATCTAAACTACATAGATCTCCAGGATTTCTTTGAGGCGCTA	982
OY	1044	CTAAAGAGGTTATTAATTTGGACAAATTAAGCCAGTTTC-----TTTATGATGAGACTCAA	1097
Dd	983	CAGAAGCATATTAATTTGGCAACAATTTCTTCGTCATCTCTCAATCTTTAGGMAACTCTGA	1042
OY	1098	GTAACACTTCAGTTTGGCCAGTACTGCCAAACATGTGGAANAATCTCCCATGTTAATGAG	1157
Dd	1043	GTACATTTGGAATATGCTCATATAGAGCAAAAGAACATATTTGATTAAGGCTGAAGATCAGA	1102
OY	1158	TTCCTGATGATGAAGGTTGCTTAAAGSTACAGAAAGGAATCTTGGATTTAAA	1212
Dd	1103	AATCCACCAAAAAAGCTCTTATTAAGGATATACGAGAGAGATTAAGCGTTTAAA	1157

```

RESULT 14
US-09-592-037-5
; Sequence 5, Application US/09592037
; Patent No. 6437115
; GENERAL INFORMATION:
; APPLICANT: Wood, Kenneth
; APPLICANT: Finer, Jeffrey
; APPLICANT: Berard, Christophe
; APPLICANT: Mak, John
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Methods of screening for modulators of
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
; FILE REFERENCE: 1014A
; CURRENT APPLICATION NUMBER: US/09/592,037
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 09/428,156
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Human

```

```

US-09-592-037-5
Query Match          1.0%;   Score 95.8;   DB 4;   Length 1542;
Best Local Similarity 53.5%;   Pred. No. 5,8e-14;
Matches 286; Conservative 0; Mismatches 222; Indels 27; Gaps 3

```

QY	684	AAAGGCTGAA	AAAAAACA	GACATTAT	TGAGAGAC	TAAATGAT	GTATC	TAGTACGCT	743
Db	644	AAAAGGGG	CAGCAAAA	AGGACAA	CTCAGCT	ACTCTGAT	GAAATGCAT	CTCTAGCT	703
QY	744	CACATACAT	ATTATTAG	ATGATTTG	TGAAGCC	GAGACAGAA	TGATCC	CACAAAATTCAG	803
Db	704	CCCCTCAG	TTTCTCTG	TTCACAT	ACATATG	AAAGA-----	AACTACGA	748	
QY	804	AGACTGTGAT	GGAGCTGT	CATGATCT	CATCTG	TAATTTGG	TAGATCTT	GGTCGACGTG	863
Db	749	TTGTATGGA	AGAAAGCT	TGTTAAAT	CGGAAGTT	GAACTTGG	TGATCTTT	GCACGAAATG	808
QY	864	AAAGAGCA	AGCCAACT	GGAGCTG	AAAGGTG	AGACTTA	AGAAAGGCT	CAACATCAAC	923
Db	809	AAAACAT	TGGCCGCT	CTGGAGCT	GTGGATTA	AGAGACCT	GGGGAAGCT	GGAATATTAATTC	868
QY	924	GCACCTTGT	TATTCCTT	GGAGAGTT	ATTAGAAG	TTTAGCGA	CGCCGCGCT	GGTGAGT	983
Db	869	AATCCCTG	TGTGACTT	TGGGAAGG	GTCTTTAC	TGCTCCCT	TGTAGAA	MAACACTCATG---	925
QY	984	TTATTAAC	TACAGAGAC	AGCAAACT	CAACGAA	TATCTC	CAAAATTC	TATGGGAGGAAATG	1043
Db	926	---TTCTT	TATCGAGAT	TTAAACTA	ACATGAA	TCTCCAGG	ATTCCTT	TGGGAGGCGGA	982
QY	1044	CTAATAAG	GTTTAAATTT	TGCAATTA	ACGCAATT	CGCATTC-----	TTTTGAT	TGAGACTCTTA	1097
Db	983	CAAACAA	ATCTATAT	TATGCAACA	AAATTTCT	CTCTCACT	CTCAATCT	TGAGAACTCTGA	1043
QY	1098	GTACACTT	CAGTTTGG	CCAGTAC	TGCCCAAC	ATGTGAA	ATACTCC	CCATGTTAAATGAG	1157
Db	1043	GTACATTG	GAATATAT	GTCTCAT	TAGAGCAAA	AGAACATAT	TGAAATTA	AGCCTGAAATGAAATCAGA	1100
QY	1158	TCTCGAT	GTATGA	AGCGTCT	CTTAAAA	AGATAC	GAAGAAATCT	TGGATTTAA	1212
Db	1103	AACGTAC	CAAAAAAG	CTCTTATTA	TAGAGAGAT	TATCGG	AGAGAT	ATAGAACCTTTAAA	1157

```

RESULT 15
US-09-724-519-7
Sequence 7, Application US/09724519
Patent No. 6414121
GENERAL INFORMATION:
APPLICANT: Wood, Kenneth
APPLICANT: Finer, Jeffrey
APPLICANT: Beraud, Christophe
APPLICANT: Mak, John
APPLICANT: Sakowicz, Roman
TITLE OF INVENTION: Methods of screening for modulators of
TITLE OF INVENTION: cell proliferation and methods of diagnosing cell
TITLE OF INVENTION: proliferation states
FILE REFERENCE: 1014A
CURRENT APPLICATION NUMBER: US/09/724,519
CURRENT FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 09/592,037
PRIOR FILING DATE: 2000-06-12
PRIOR APPLICATION NUMBER: 09/428,156
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1728
TYPE: DNA
ORGANISM: Human
US-09-724-519-7
Query Match 1.0%, Score 95.8, DB 4, Length 1728;
Best Local Similarity 53.5%, Pred. No. 6, 1e-14;

```

Matches 286; Conservative 0; Mismatches 222; Indels 27; Gaps 3;

```
OY 684 AAAGGGTGAAGAAACAGACATTATGAGAGAGATAAATGATGATAGTACTGTT 743
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 644 AAAGGGGGCGAGAAAGAGACACTGACCTCTGATGATGATGATGATGATG 703
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 744 CACATACATATTTAGATGTTGTTGAAGCCGAGACAGACAAATGATCCACAAATTCAG 803
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 704 CCCACTCAGTTTCTCTGTTACATATACATGAAAGA-----AACTACGA 748
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 804 AGAAGTGTGATGAGAGCTGTGATGATCTGATGATGATGATGATGATGATG 863
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 749 TTGATGGAGAGAGCTTGTAAATCGAAAGTTGAAGTGTGATGATGATGATG 808
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 864 AAAGAGCAAGCCAAACTGAGCTGAGGTGAGACTTGAAGAGAGGCTGCAACATCAAC 923
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 809 AAACATTTGGCCCTTGTGAGCTGTGATGATGAGAGCTGAGAGCTGGAATATTAATC 868
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 924 GCAGCTTGTATCTCTGGAGAGTTTAAAGAGCTTGAAGAGGAGGAGGCTGATGAT 983
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 869 AATCCCTGTGACTTTGGAGAGGTCATTTACTGCCCTTGTAGAAAGAACCTCATG--- 925
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 984 TTATTAATCTACAGAGACAGCAAACTCACAGAAATTCCTCAAAATTCATTTGGAGAGAAATG 1043
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 ---TTCCTTATCGAGATCTAAACTACTAGATCTCCAGAGATTCCTTGGAGGGCTA 982
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1044 CTAAAGGTTTAAATTTGCACATTTAGCGCAGTTTC-----TTTGTAGAGACTTAA 1097
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 983 CAAGAACATCTATTAATTTGACAACTTCTCCGCAATCTCTCAATCTTGAGGAAACTCTGA 1042
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1098 GTACACTTCAGTTTGGCAGTACTGCCAACAATGTGAGAAATCTCCCATGTTATGAGG 1157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1043 GTACATTTGGAATGCTCATAGAGCAAGAACATATTTGATTAAGCCTGAGAGTGAATCAGA 1102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1158 TCCCTGAGTGTGAAGGTTGCTAAAGGTACAGAAAGGAATCTTGATTTAA 1212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1103 AACTCACCAGAAAGCTCTTATTAAGAGATTAAGAGAGATAGAACGTTTAA 1157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: May 23, 2003, 01:06:49
Job time : 482 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 00:54:39 ; Search time 1088 Seconds
(without alignments)
11682.693 Million cell updates/sec

Title: US-09-150-867-2
Perfect score: 9626
Sequence: 1 gaattccgagtcgtagtag.....ttaaaaaaaaaaagcaattc 9626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PC1_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PC105_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620.4	6.4	8493	US-10-071-766-51	Sequence 51, App1
2	615.2	6.4	7992	US-09-893-519A-140	Sequence 140, App
3	206.6	2.1	581	US-09-833-790-351	Sequence 351, App
4	138.2	1.4	501	US-09-833-790-167	Sequence 167, App
5	117.6	1.2	3366	US-09-938-842A-2651	Sequence 2651, App
6	95.8	1.0	4858	US-09-954-456-733	Sequence 733, App
7	90	0.9	29793	US-09-973-451-38	Sequence 38, App1
8	85.6	0.9	501	US-09-833-790-329	Sequence 329, App
9	81.2	0.8	6586	US-09-954-456-1115	Sequence 1115, App
10	81.2	0.8	6586	US-09-954-456-1790	Sequence 1790, App
11	73.8	0.8	3348	US-09-954-456-90	Sequence 90, App1
12	73.8	0.8	3348	US-09-954-456-727	Sequence 727, App
13	68.2	0.7	5361	US-10-071-766-80	Sequence 80, App1
14	68	0.7	5361	US-09-742-096-2	Sequence 2, App1
15	68	0.7	6152	US-09-742-096-1	Sequence 1, App1
16	67.8	0.7	5848	US-10-044-090-299	Sequence 299, App
17	67.6	0.7	3930	US-09-847-874A-2	Sequence 2, App1
18	59	0.6	9539	US-10-239-676-52	Sequence 52, App1
19	58.4	0.6	1002	US-09-879-536-442	Sequence 442, App

20	58.2	0.6	498	US-09-854-133-345	Sequence 345, App
21	58.2	0.6	498	US-09-738-973-345	Sequence 345, App
22	58	0.6	529	US-09-983-965-2109	Sequence 2109, App
23	57.4	0.6	446	US-09-960-352-3400	Sequence 3400, App
24	55.6	0.6	4267	US-10-001-843-45	Sequence 45, App1
25	54.4	0.6	451	US-09-960-352-10262	Sequence 10262, App
26	53.2	0.6	728	US-09-764-898-109	Sequence 109, App
27	53.2	0.6	1488	US-09-764-898-37	Sequence 37, App1
28	53	0.6	516	US-09-960-352-5785	Sequence 5785, App
29	52.8	0.5	589	US-09-864-721-4741	Sequence 14741, App
30	50.6	0.5	12405	US-10-239-676-35	Sequence 35, App1
31	50.4	0.5	554	US-10-101-487-69	Sequence 69, App1
32	50.4	0.5	554	US-10-101-487-106	Sequence 106, App1
33	50.4	0.5	1014	US-09-883-096-3	Sequence 3, App1
34	50.4	0.5	1152	US-09-883-096-6	Sequence 6, App1
35	50.4	0.5	4108	US-09-883-096-1	Sequence 1, App1
36	50.2	0.5	522	US-10-101-487-71	Sequence 71, App1
37	50.2	0.5	530	US-10-101-487-73	Sequence 73, App1
38	49.4	0.5	7657	US-10-239-676-185	Sequence 185, App1
39	49	0.5	17848	US-09-938-842A-646	Sequence 646, App
40	48.8	0.5	2241	US-10-171-311-7	Sequence 7, App1
41	48.6	0.5	12313	US-10-171-311-7	Sequence 7, App1
42	48.6	0.5	12337	US-10-171-311-5	Sequence 5, App1
43	48.6	0.5	12438	US-10-171-311-3	Sequence 3, App1
44	48.6	0.5	12462	US-10-171-311-1	Sequence 1, App1
45	48.4	0.5	431	US-09-960-352-5558	Sequence 5558, App

ALIGNMENTS

RESULT 1	US-10-071-766-51
Sequence 51, Application US/10071766	
Publication No. US20020192678A1	
GENERAL INFORMATION:	
APPLICANT: Huel-Mei Chen	
TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE	
FILE REFERENCE: PA-0043 US	
CURRENT APPLICATION NUMBER: US/10/071,766	
CURRENT FILING DATE: 2002-02-07	
NUMBER OF SEQ ID NOS: 144	
SOFTWARE: PERL Program	
SEQ ID NO 51	
LENGTH: 8493	
TYPE: DNA	
ORGANISM: Homo sapiens	
FEATURE:	
NAME/KEY: misc_feature	
OTHER INFORMATION: Incyte ID No. US20020192678A1 441283.3	
US-10-071-766-51	
Query Match	6.4%; Score 620.4; DB 9; Length 8493;
Best Local Similarity	65.5%; Pred. No. 9e-132;
Matches 1001; Conservative	0; Mismatches 491; Indels 36; Gaps 5;
QY	69 GCGGCCAAGAGGACAAAGTACAGACAGACAGGAGCGGTGCGTACCGATTTCAC 128
DB	17 GGGGCTGTGAGCCCTGAAGTGCAGGCGGCGGAGGCTGCGCATTTTGTTGGACAG 76
QY	129 TATCGGTCTCAAAATGTCCGAGGAGATGATCAATTAATGTGTGTGAGGTTCGGCCGC 188
DB	77 TTCAGCTATATGATGATGCGGAGGAAGAGCGGTGCGTTCGCGTGCAGTGGCGCCGC 136
QY	189 TTAATACAGAGACAA-----GGGATCAAGCCCACTGCAATGGAAGGCTGCAACA 242
DB	137 TCAACAGCAGAGATCACTTGGGAAACGCCCAATTTACTGGAACAGTCAACA 196
QY	243 AACACATTTCCCAAGTGTGAGGAGCAAGTCTTCAATTCATGATGATTAATTC 302
DB	197 ATGCTATTATTAAGTTGATGAGTAAGTAATCCCTTCATTTGATCGTCTTCATGTA 256
QY	303 ACGAATCAACAAGTCAAAATTTACCAAGAAATAGCAGTACTATATCATCATGCTTTC 362

```

Db 257 ATGAACATACCAAAATGCTATGATGAAGAAATGAGCACCACATCATGATCTGCGCATAC 316
QY 363 AGGATATATATGGGCAATATTTGGATACGAGACATCTTCGAGGAGACATACAA 422
Db 317 AAGGCTACAAATGATCTATATTTGGCTATGAGACAGCTGCTTGAGAAAAACATATATCA 376
QY 423 TGATGGGAACCAAAATTCATTTGGGCAATATATCCCAAGCCATACAGGAGTTTAA 482
Db 377 TGATGGGTTTCAAGATCATTTGGGAGTTATACCAAGGCAATTCATGACATTTCCAA 436
QY 483 TTATTCAGGATACCCGAGAGAGAGTTCTTCAAGAGTTCTTATATGAGATTTCA 542
Db 437 AAATTAAGAAATTCCTGATAGGAAATTCCTTACAGTATCTTACATGAAATATACA 496
QY 543 ATGAACATGTCGAAAGACCTATGCTGATGACAGAAAGAAAGACCCCTGGAAATGCG 602
Db 497 ATGAACCATATACAGATTTACTGCTGCACTCAAAAATTAACCTTATTTTGAG 556
QY 603 AGGATTTATAGAAAGCTGATATGCTGACCTGACCTGAGAACTTGAATGCTCTG 662
Db 557 AAGATGTCAAATGGAATGTATGTCTGATCTACAGAAAGATGTATATACATCAG 616
QY 663 AACATGTAATACAGTGGATCAAAAGGTTGAAACAAACGATATATGAGAGCTAAA 722
Db 617 AATAGCTTTGAATGATATGATACAAAGGAGAAAGACGAGCATATAGAGAAACAAA 676
QY 723 TGAATGATCATATAGTATGCTTACATCAATATTTAGATGATTTGAAACCGAGACA 782
Db 677 TGAATCAAGAGAGCTGCTTCTCATACATCTTAGATGATTTTGGAAAGCAGAGAGA 736
QY 783 GAATGATCCCAAAATTCAGAGACTGTGATGAGAGCTGTCAAGTATCTCACTGAA 842
Db 737 AGGCTGAGACCTTCTAAAT-----TGATGAGATCTGTTAAGGTATCCCATTTGA 787
QY 843 TGTAGATCTTGTGCTGAGTGAAGAGACCAAGCCAACTGGAGCTGAGAGTGTGAGACT 902
Db 788 TGTGATATCTTGCAGGAGAGTGAAGAGAGCTGCTCAAAAGGCGTGGAGTGGGCTCA 847
QY 903 AGGAAGCTCAGACATCAACCGAGCTGTTTATCTTGAAGAGTATTTAGAGCTTA 962
Db 848 AGGAAGCTGATATATATATATGAAAGCTTATTTTGGGCAAGTATCAAGAACTTA 907
QY 963 GCGAGCGGCGAGCTGCTGATATATATATACAGAGACAGCAAACTCAGCAATTTCC 1022
Db 908 GTGATGACAAAGTGGGTTGCTTAAATATATGAGATACAGGATTAACGAAATCTTC 967
QY 1023 AAAATTCATGAGAGAAATGCTAAACGGTATATATTTGACAAATATCCGAGTTTCT 1082
Db 968 AGAATTCCTTGGGAGGAAATCCAAAGACAGTATTTATCTGACAAATATCTCCGATCT 1027
QY 1083 TTGATGAGACTTATAGTACACTTCAGTTTCCAGTACTGCGCAAAACATGTGAGAA 1142
Db 1028 TTGATGAACTCTTACTGCTCTCCAGTTTCCAGTACTGCTAAATATATGAGAAAT 1087
QY 1143 CCCATGTTATGAGAGCTGATGAGAGCTGTGCAAAAGGCTACAGAAAGAAATCT 1202
Db 1088 CTTATGTTATGAGATGATCACTGATGAGAGCTCTCGAAAGAGTATGAGAAAGATA 1147
QY 1203 TGTATTTAAAGAAACATTTAGAGATTCATGCTGTAAGCAAAAGCTCAGCA 1262
Db 1148 TGTATCTTAAAGAAACATTTAGAG-----GAGTTTCTTATGAGAGCGGGCTCAGCA 1201
QY 1263 TGTCTTAAAGAGAGCATACACAGCTTGTAGTAAATCAAAACAGTACACAGAGAG 1322
Db 1202 TGTGAAAGAGCAATTTGGCCCACTTTTGGAAAGAAAGATTTGCTTCAGAAAGTACA 1261
QY 1323 AAGATGAAATATGCACTGCAAAATATTTGTG-----CTTCATCCCAAGATCTCAAC 1379
Db 1262 ATGAGAAATTTGAAATCTTACAGAGATGCTGTGACCTCTTCTTCCCTCAGCTTCAAC 1321
QY 1380 AGGACCAAAAGGCTCAAGCAAGAAAGATTTAGTGGGCGCAGAGAAATTCAGAAATA 1439

```

```

Db 1322 AGAATTAAGAGCTTAAAGAAACGAGAGTACTGTGCTGCTTGGCAAAATTAACAAA 1381
QY 1440 GTTACATGCTTCTGGGTTTCTGACTTGTATATGCTATCAATTTACCTGGCAATTTA 1499
Db 1382 TGAGAA-----CTCAACATGATGAGATCAATTAATATTCACAAAAATTA 1429
QY 1500 GCAAGAGGCAAGGCTCTGACATGCTTCATTTCCAGAAATGATGACTGTGTTGA 1559
Db 1430 CAACAAAACACATACACTTCTTATTAATTTATACAGAAATTTGATGATCTGTCTGT 1489
QY 1560 CAGAGTTTCTGATTTGATGAGCGCCT 1587
Db 1490 CAGAGCTGATGTTTTCAGTAACTCT 1517

RESULT 2
US-09-893-519A-140
; Sequence 140, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thimara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAO, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/16548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/215005.1
; DATABASE ENTRY DATE: 1993-01-10
; RELEVANT RESIDUES: (1)..(7992)
US-09-893-519A-140

Query Match 6.4%; Score 615.2; DB 9; Length 7992;
Best Local Similarity 66.7%; Pred. No. 1.4e-130;
Matches 970; Conservative 0; Mismatches 448; Indels 36; Gaps 5;

```

```

Db 181 AATGTATGAAAGATATACAGACCAATCATCGATTCGCCATACAAAGCTACATGCT 240
Qy 377 ACAATATTTGCTATGAGAGACAGACATCTTGAGCAGACAGCTACACATGATGGACACCA 436
Db 241 ACTATATTTGCTATGAGAGACAGACATCTTGAGCAGACAGCTACACATGATGGACACCA 300
Qy 437 AATTCATTTGGGATATATACCCCAAGCATACAGAAAGTTTAAATTTATTCAGAGATA 496
Db 301 GATCATTTGGGATATATACCCCAAGCATATGACATTTCCAAAAATTAAGAAAGTTT 360
Qy 497 CCGAAGAGAGAGATCTTCTTAAGAGTTTCTTAATGAGATTTACATGAAATGCTGAAA 556
Db 361 CCTGATAGGAGAAATTTCTTACGTGATCTTACATGAAATATACAAATGAAACCATTACA 420
Qy 557 GACCTACTGTGTGATGACAGAAAGAAAGCCCTTGAAATTCGCGAGATTTTAAATAGA 616
Db 421 GATTTACTCTGTGAGACCTCAAAAAATGAAACCTTTAATTTATCGAAGAAATGTCATPAG 480
Qy 617 AACGTGATGTGCTGACCTGACCTGAGAAACCTTGTAAATGTTCTGTAACATGTAATACAG 676
Db 481 AATGTATGATGTGCTGATCTCAGAGAGAAATGTAATACATCAGAAATGCTTTGAAA 540
Qy 677 TGGATCAAAAAAGGATGAAAAAAGACATATATGAGAGACTAAATGATGATCATAGT 736
Db 541 TGGATTAACAAAGGAGAAAAAGAGAGGCAATATGAGAAACAAAAATGATCAAGAAAGC 600
Qy 737 ACTGCTTCAATCAATATTTAGATGATTTGTAAGAACCCGAGACAGAAATGATCCACA 796
Db 601 ACTGCTTCAATCAATATTTAGATGATTTGTAAGAACCCGAGAGAAAGGCTGACCTTCT 660
Qy 797 AATTCAGAAAGATCTGATGAGCTGATGATGATCATCTTAATTTGTTAGATCTTGT 856
Db 661 AAT-----TGTAAAGATCTGTTAAAGTATCCATTTGAATTTGTTGTTGTA 711
Qy 857 GGCAGTGAAGAGCAAGCCAACTGAGCTGAGGTGAGACTTAAAGAGAGCTGCAAC 916
Db 712 GGCAGTGAAGAGAGCTGCTCAAAAGAGGCTGAGGTGCTGAGGCTCAAGAAAGCTGTAAAT 771
Qy 917 ATCAACCGCAGCTGTTTATCTTGTGACAGGTTTATTAAGAGCTTAAAGCGCCAGCT 976
Db 772 AATTAATCGAAGCTTATTTATTTGAGACAAAGTATCAAGAACTTATGATGACAAAGT 831
Qy 977 GGTGATTTATTAATCAAGAGACAGCAAACTCACCAGATTTCTCCAAATTCATTTGAGGA 1036
Db 832 GGTGATTTATTAATCAAGAGATGAGATTAACACGATTTCTCAGATTTCTGAGGA 891
Qy 1037 GGAATGCTAAAAAGCTTATATTTGACAAATTAAGCCAGTTTCTTTGATGAGACTTA 1096
Db 892 GGAATTCAAAAGACAGATATATGACAAATTAAGCCAGTTTCTTTGATGAGAACTTT 951
Qy 1097 AGTACACTTCACTTGGCCAGTACTGCAAAACATGAGAAATTAATCTCCATGTAATGAG 1156
Db 952 ACTGCTTCCACTTGGCCAGTACTGCTAAATATGAAATACTCTTATGTAATGAG 1011
Qy 1157 GTCCTGATGATGAGAGCGTGTCTTAAAAAGGTACAGAAAGAAATCTGATTTAAAGAA 1216
Db 1012 GATCACTGATGAGAGCTCTCTGAAAAGGTATAGAAAAGAAATATGATCTTTAAAAA 1071
Qy 1217 CAATTAGAAATTTAGAGTCACTGCTGAACAAAAGCTCAAGCAATGCTTAAAGAGAG 1276
Db 1072 CAATTAGAG-----GAGGTTCTTTAGAGAGCGGCTCAAGCAATGAGAAAAAGCCAA 1125
Qy 1277 CATACACAGTTGCTAGCTGAATCAACACATACAGAAAGAGAGAGATGAAATATG 1336
Db 1126 TTGGCCCAACTTTTGAAGAAAGATTTGTCAGAAAAGTACAGAAATGAGAAATTTGAA 1185
Qy 1337 CACTTGACAAATATTTGTTGT---TGCTTCATCCCAAGATTCACAGAGACCAAAAGGCTC 1393
Db 1186 AACTTAACACAGATCTGCTGACCTCTTCTTCTCCTCAGCTTCAACAGAGAAATTAAGGCT 1245
Qy 1394 AAACGAAAAAGAGATTAAGTGGGCGCAGAAAAATCCAAATAGTTTACATGCTTCT 1453

```

```

Db 1246 AAAAGAAAAAGAGATTAAGTACTTGCTGCTTGGCAAAATTAACAAATTAAGAAAGCTCAAC 1305
Qy 1454 GGTTGTTGACTTTATATGCTATGCTATGCTATGCTGCAATTTTACAAAGAGCAAG 1513
Db 1306 TATGAGATCAATTTATATATACCA-----AATTAACAAACAAACAAACAT 1353
Qy 1514 TTCTGACATGCTTCAATTTCCAGAAATTTGATGACTGCTGTTGATGAGAGTTTCTGAT 1573
Db 1354 AACCTTCTATTAATTTATATATAGCAAAATTTGATGAAATCTGCTGCTGACAGTGTAT 1413
Qy 1574 TTTGATGACGCCCT 1587
Db 1414 TTCAGTAACATCT 1427

RESULT 3
US-09-833-790-351
; Sequence 351, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tonglong
; APPLICANT: Secretist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 351
; LENGTH: 581
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-790-351

Query Match 2.1%; Score 206.6; DB 10; Length 581;
Best Local Similarity 68.4%; Pred. No. 2.6e-37;
Matches 333; Conservative 0; Mismatches 144; Indels 10; Gaps 3;

Qy 955 GAAGCTTACGAGCGCGGAGCTGCTGATTTAACTACAGAGACGAACCTCACAG 1014
Db 72 GAAGCTTACGAGCGAGAGTGTGCTTCAATTAATTCAGATGAGCAAGTTAAACAG 131
Qy 1015 AATTCCTCAAAATTCATTTGGAGAAATGCTAAACGTTAATTAATTTGCAATTAAGCC 1074
Db 132 AATTCCTCAAAATTCCTTTGGAGAAATGCAAAAGACGATATATCTGCAATTAATCTCC 191
Qy 1075 AGTTCTTTTATGAGAGCTCTAATGATCACTTCACTTGTCCAGTACTGCCAAACATGTAG 1134
Db 192 AGTATCTTTTATGAGAAACCTTACTCTCCAGTTGCTCAGTACTGCTAATATATGAA 251
Qy 1135 AATATCTCCCATGTTAATGAGGCTGCTGATGATGAAGGTTGCTTAAAGATGAGAA 1194
Db 252 GAATATCTCTTATGTTAATGAGGATCAACTGATGAGGCTCTCTGAAAAGATATGAAA 311
Qy 1195 GGAATCTGGAATTTTAAAGAAACATTTAGAGATTTAGATCATCTCTGAAAACAAAGC 1254
Db 312 AGAATATTAATGATCTTAAAAAAACAATTAGAG-----GAGGTTTCTTTAGAGCGGCGC 365
Qy 1255 TCAAGCATGCTTAAAGAGAGATACACAGTGTGCTAGTGAATCAAAACATACACA 1314
Db 366 TCAGGCAATGAGAAAGAACCAATTTGCCCCACTT-TTGGAAAGAAAAAGATTTCTTCAGAA 424
Qy 1315 AGAGAGAGAGATTAATGATGACCTTGACAAATATGTTGT---TGCTTCATCCCAAGA 1371
Db 425 AGTACGAATGAGAAATTTGAAACCTTAACAGGATGCTGTGACCTCTTCTCCCTCAC 484
Qy 1372 ATCTCAACAGACCAAAAGGCTCAACGAAAGAGATTTAGTGGGCGCCAGGAAAT 1431

```

```

: APPLICANT: Harper, Jeff
: APPLICANT: Krebs, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 2651
: LENGTH: 3366
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: US-09-938-842A-2651

```

7

QY 1185 GGTACAGAAAGAAATCTGATTTAAAGAAACATTAGA 1224
DB 1517 AATATCAAGAGAAATCTCAACCTTCAACTAGAACTTGA 1556

RESULT 6
US-09-954-456-733

; Sequence 733 Application US/09954456
; Patent No. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; CURRENT FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,638

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,863

; PRIOR FILING DATE: 2000-09-27

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 733

; LENGTH: 4838

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-456-733

Query Match 1.0%; Score 95.8; DB 10; Length 4858;
Best Local Similarity 53.5%; Pred. No. 3e-11;

Matches 286; Conservative 0; Mismatches 222; Indels 27; Gaps 3;

QY 684 AAAAGGCTGAAAAAACAACATTTATGAGAGACTAAATGAATGATCATAGTCGTT 743
DB 734 AAAAGGGGACGAAAAAGACACATGCACTGCTGATGAATGATCATAGTCGTT 793
QY 744 CACATACAAATTTAGAAATGATTTGAAAGCCGAGACAGAAATGATCCCAAAATTCAG 803
DB 794 CCACACTGATTTCTCTGTTACAAATATATGAAAGA-----AACTACGA 838
QY 804 AAAACTGTATGAGAGCTGTCATGCTATCTCAGCTGTAATTTGGATGATCTTGGCGAGTG 863
DB 839 TTGATGAGAAAGAGCTTTGTTAAATGGAAGTTGAACCTGTTGATCTTGCAGGAAGTG 898
QY 864 AAAGAGACGCAAACTGAGCTGAGTGTGAGTGAAGTGAAGGCTGCAACATCAACC 923
DB 899 AAAACATTTGGCGCTTGTGAGCTGTGATTAAGAGAGCTGGGAAAGCTGCAAAATTAATC 958
QY 924 GCAGCTTTTATCTTGGACAGGTTATTAAAGGCTTAAGCAGCGGCGCTGTGGAT 983
DB 959 AATCCCTGTGACTTTGGGAGGCTATTAAGTCCCTGTAGAAAGAACACCTCATG--- 1015
QY 984 TTATAACACAGAGACAGCAAACTACAGAAATTTCTCAAAATTTATGCGGAGAAAG 1043
DB 1016 ---TTCTTATGAGAAATCTAAACTAAGTAAATCTCCAGGATTTCTTGGAGGCGTA 1072

QY 1044 CTAAGAGCTTATATTTGCAATTTAGCCAGTTTC-----TTTGATGAGACTCTAA 1097
DB 1073 CAAGACATCTATATATTTGCAACAAATTTCTCTGATCTTCATCTTTGAGGAAGTCTGA 1132
QY 1098 GTACACTTTCAGTTTCCGAGTACTGCCAAGATGAGAAATTAATCTCCCATGTTAAAGAG 1157
DB 1133 GTACATTTGAAATATGCTCATATGAGCAAGAAATATTAATTAAGCCCTGAAGTGAATCAGA 1192
QY 1158 TCCTGATGATGAAGCGTCTCTTAAGAAAGTACAGAAAGAAATCTTGATTTAA 1212
DB 1193 AACTCACCAAAAAAGCTTTATTAAGGATATACGAGAGATGAAGAGCTTTAA 1247

RESULT 7

US-09-973-451-38/c

; Sequence 38, Application US/09973451

; Patent No. US2002013238A1

; GENERAL INFORMATION:

; APPLICANT: JACOBSON, MYRON K.

; APPLICANT: JACOBSON, Elaine L.

; APPLICANT: AM, Jean-Christophe

; APPLICANT: LIN, Winston

; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE

; TITLE OF INVENTION: (PARG) ENZYMES, THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREAC

; TITLE OF INVENTION: THEREWITH

; FILE REFERENCE: N1AD 201

; CURRENT APPLICATION NUMBER: US/09/973,451

; CURRENT FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: US/09/302,812

; PRIOR FILING DATE: 1999-04-30

; PRIOR APPLICATION NUMBER: 60/083,768

; PRIOR FILING DATE: 1998-05-01

; NUMBER OF SEQ ID NOS: 38

; SEQ ID NO 38

; LENGTH: 29793

; TYPE: DNA

; ORGANISM: Caenorhabditis elegans

; FEATURE:

US-09-973-451-38

Query Match 0.9%; Score 90; DB 10; Length 29793;
Best Local Similarity 50.8%; Pred. No. 2e-09;

Matches 309; Conservative 0; Mismatches 275; Indels 24; Gaps 3;

QY 716 ACTAAATGATATGATCATAGTATGCTGTCATACATCAATTTATGATGATTTGAAAGC 775
DB 17791 ACCAACATGATGATGATGATGCTGCTGCTCCAGCTATCTTTATCATTTAGTGAATG- 17733
QY 776 CGAGACGAAATGATCCCAAAATTCAGAGAGTGTGATGAGTGTGATGATGATGATGATG 835
DB 17732 -----TTCCCAATTTGTTGAGAGCGGAGAAAGTATTAATTAAGTTGCTGA 17687
QY 836 TTGAATTTGATGATCTTCTGTCAGTGAAGAGCAAGCCAACTGAGCTGAAGTGTG 895
DB 17686 CTAATCTGTATGATTTAGCTGTGTCAGAACGGCAATCGAAACAGAGCAAGCTGGGAA 17627
QY 896 AGACTTAAGAGAGCTGCAACATCAACCGACGCTGTTATCTTGGACAGTTATTAAG 955
DB 17626 AGATTTAAAGAACTACGAAATCAATCTTCAATTCGGCTGTTGAAATGTAATCAGT 17567
QY 956 AACCTTAGCAGCGGCGGCTGATTTATTAACACAGAGCAAGCAACTCACAGA 1015
DB 17566 GCATTTGATGATCGAATC---TGCCCAATTTCTTATGATGATTAATCAAACTGATCGA 17510
QY 1016 ATTCTCCAAATTCATTTGAGAGAAATGCTAAACGGTTAATTTGCACAATTAAGCCA 1075
DB 17509 CTCTCAAGATTTCTCGGTGGAATTAAGACAGTTATGTTGCTGTTGAGACT 17450
QY 1076 G-----TTTCTTTGATGAGACTCTTAAGTACCTTCACTTCACTGAGTACTGCCAAAGAT 1129
DB 17449 GCAAGTTACAATTTTGAAGAGACATTTGATCATTAAGATATGCAATGAGCTAAAMAC 17390

Db	1707	TCATGAGAAAGCTAGACAGCACACCTCTCTAGAAAGACCTTGGAAAGCTTCAATACCTGTC	1766
Qy	381	TATTTGCATACGACGACAGCATTTCTACAGGCACAGCTACACAAATGATGGGACACCAAAATT	440
Db	1767	TTTTTGGTTATGGGACACACTGGCTCTGGAAAAATCATATACGATGATGGGATTTACTGTAAG	1826
Qy	441	CATTGGGCATATATACCCGACAGCCATACAGGAAGTTTTTAAATTTATTCAGAGATACCGA	500
Db	1827	AACGAGAAATATATCCAAAGATTTTGTGAGATCTTTTTCTCAAGTAGCCAGAAAAACAA	1886
Qy	501	ACAGAGAGTT-----TCCTTAAAGATTTCTTATATGAGATTTTACAATGAAGCTGGA	554
Db	1887	CCCAAGAGGCTACGCTATACCATTTGAATGAGCTTCTTGAATATATATGAAAAAATTC	1946
Qy	555	AAGACACTACCTGTG-----ATGACAGAAAGAAAGAGCCCTTGGAAATTCGG	602
Db	1947	ACGACCTCTTGCTTTGTAAGAATGAAAAATGGGACAGAGAAAGAACCCACTGACAGATGACGG	2006
Qy	603	AGGATTTTAAATAGAAACGTATGTATGCTGACCTGACTGAACAACTTGTATATGTTCTCTG	662
Db	2007	AACATCTGTTTATGAGACCATATGTGTAAGACATGTCATGAACATTTGCACTGTTTACG	2066
Qy	663	AACATCTGAATACAGTGGATCAAAAAAGGCTGAAAAAACAGACATTATGGAGAGACTAAA	722
Db	2067	CTGATATACCAAGATTGCTGAATTTGGGAAATTAACAAAGAGACTACTGCTGCTACTGGTA	2126

```

Db 2127 TGAATGATTAAGTAAGTTCGCCGATCTCATCTCAGTTTTCACCTCGTGATGACCCAGACCCAGA 2186
Oy 783 GAATGATCCACAAATTCAGAGAACTGTGATGAGCTGTCTCATGGTATCTCACTTGAAT 842
Db 2187 CAGAAATTTGTGGAGGGGAGAACACGATCACAGA-----ATAACAGTCCAAATTAACC 2240
Oy 843 TGTGATCTTCTCTGSCAGTCGAAAGACGAACTGGAGCTGAGAGTGTGAGACTTA 902
Db 2241 TAATGATCTGCGACGCACTGAGCCCTGCTCTACGGCTCACACTAATGAGATTCGACTTA 2300
Oy 903 AGGAAGCTGCACAACTCAACCGCAGCTTGTATTCCTTTGGACAGGTTATTAAGAACTTA 962
Db 2301 AGGAAGGTGTGAGTATTAATAGTCTTGCCTTAACCTTGGGAAAGTATATCTGCACCTT 2360
Oy 963 GCGA-----CGCCAGCGCTGTGATTTATTAATACAGACAGCAACACCTCACCAGA 1016
Db 2361 CGGAACAGCAAAACCAAGAGAGTGTATTTATCTTATTCGTGAATCTGTCTTACATGGC 2420
Oy 1017 TTCTCCAAATTCATTTGGGAGGAGAAATGCTAAACGGTATTAATTTGCAAAATTACGCCAG 1076
Db 2421 TGTTAAGAAGAAAGTGTGGGTGGAATTCAAAACATGCATGATTTGCTACGATTAAGTCCG 2480
Oy 1077 TTTC-----TTTGTATGAGACTCTAAGTACACTTCAGTTGGCCAGTACTGCCAAACATG 1130
Db 2481 CTGCACGACAACTAATMAAACAATATACACACTTGAATATCTTAACCAACGCCCTTTAA 2540
Oy 1131 TGAGAAATATCTCCCATGTTAATGAGTCTGTGATGATGAGA 1170
Db 2541 TAGTCACATTTGCTTAAAGTAAATGAAGATGAGAACCTTA 2580

RESULT 11
US-09-954-456-90
Sequence 90, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
PRIORITY FILING DATE: 2001-09-18
PRIORITY FILING DATE: US/60/233,617
PRIORITY FILING DATE: 2000-09-18
PRIORITY APPLICATION NUMBER: US/60/234,052

```

PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 90
LENGTH: 3348
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-90

Query Match 0.8%; Score 73.8; DB 10; Length 3348;
Best Local Similarity 50.9%; Pred. No. 2.7e-06;
Matches 202; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

OY 597 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 656
DB 815 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 874
OY 657 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 716
DB 875 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 934
OY 717 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 773
DB 935 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 994
OY 935 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 833
DB 774 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1054
OY 995 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 893
DB 834 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1114
OY 1055 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 953
DB 894 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1174
OY 1115 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 990
DB 954 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1211
OY 1211 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1174

RESULT 12
US-09-954-456-727
Sequence 727, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
FILE REFERENCE: 689290-76
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20

PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 727
LENGTH: 3348
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-727

Query Match 0.8%; Score 73.8; DB 10; Length 3348;
Best Local Similarity 50.9%; Pred. No. 2.7e-06;
Matches 202; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

OY 597 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 656
DB 815 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 874
OY 657 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 716
DB 875 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 934
OY 717 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 773
DB 935 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 994
OY 935 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 833
DB 774 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1054
OY 995 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 893
DB 834 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1114
OY 1055 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 953
DB 894 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1174
OY 1115 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 990
DB 954 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1211
OY 1211 TTCCGAGAGATTTAATAGAAAGTGTATGCTGACCTGACCTGAAGAAGCTTGAATG 1174

RESULT 13
US-10-071-766-80
Sequence 80, Application US/10071766
Publication No. US20020192678A1
GENERAL INFORMATION:
APPLICANT: Huel-Mei Chen
TITLE OF INVENTION: GENES EXPRESSED IN SENESECE
FILE REFERENCE: PA-0043 US
CURRENT FILING DATE: 2002-02-07
PRIOR APPLICATION NUMBER: US/10/071,766
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PERL Program
SEQ ID NO 80
LENGTH: 3376
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Inbyte ID No. US20020192678A1 319885.11
US-10-071-766-80

Query Match 0.7%; Score 68.2; DB 9; Length 3376;
Best Local Similarity 50.7%; Pred. No. 5.3e-05;
Matches 191; Conservative 0; Mismatches 183; Indels 3; Gaps 1;

QY 617 AACGCTGATGTTGCTGACCTGACGAACTGTAATGTTCCGTAACATGTAATACG 676
DB 959 AACAGTGTGTTGCGAGATGACAGAGTTGAGTAATCAATCACTACGAGGAGCTTTGAA 1018
QY 677 TGGATCAAAAAGGCGTAAAAACAGACATTTATGAGAGACTAAATGATGATAGT 736
DB 1019 GTTTCTTGAGAGGCGCAAGAAAGAGCGTATTTGCTAATACCATTTTGATGTCGATCC 1078
QY 737 AGTCGTTACATCAATATTTTGAATGA---TTGTTGAAGCCGACAGACAAATGATCC 793
DB 1079 AGCGTTCCATAGCGGTGTCAACATTAATAGTTCAAGGCTCCCTTGATGCAATGGA 1138
QY 794 ACAATTTAGAGAACTGATGAGAGCTGTCATGTCATCTCACTTAATTGGTAGATCTT 853
DB 1139 GACATGTCTTACAGGAAAAAGAACAAATCACTATAGCTGTCTCTGTTAGATCTT 1198
QY 854 GCTGCACTGAAAGAGCAAGCCAACTGAGCTGAGAGTGTGAGACTTAAGGAGGCTGC 913
DB 1199 GCTGAAATGTAAAGAACTACACCGAGACAGAGCAAGAGAAAGATTTAGCTGATGCT 1258
QY 914 AACATCAACCCGACGCTTTTATCTTGTGACAGGTTATTAAGACCTTAGCGAGCCAG 973
DB 1259 AATATTAATCACTCACTAATGACGCTAAGAACATGTATGATGCTTAAGAGAACCA 1318
QY 974 GCTGGTGAATTTATTA 990
DB 1319 ATGTAAGAACTAACAA 1335

RESULT 14
US-09-742-096-2
Sequence 2, Application US/09742096
Patent No. US20020155441A1
GENERAL INFORMATION:
APPLICANT: DROUHE, PIERRE
APPLICANT: DABERRES, PIERRE
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
FILE REFERENCE: 2007730SDIV
CURRENT APPLICATION NUMBER: US/09/742.096
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 08/973,642
PRIOR FILING DATE: 1998-02-06
PRIOR APPLICATION NUMBER: PCT/FR96/00894
PRIOR FILING DATE: 1996-06-12
PRIOR APPLICATION NUMBER: FR 95/07007
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 5361
TYPE: DNA
ORGANISM: P. falciparum
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(5361)
OTHER INFORMATION:
US-09-742-096-2

Query Match 0.7%; Score 68; DB 9; Length 5361;
Best Local Similarity 42.1%; Pred. No. 7.8e-05;
Matches 660; Conservative 0; Mismatches 890; Indels 18; Gaps 4;
QY 3689 AGATATGAAAGTGTGATCTCCAGAACCAAGCAAGCTTGGAAGAGAGAGGTTAA 3748
DB 4510 GTTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4569

DB 3445 ATTATTAAGAAATGGAAGAAATATTTAGATGATTAAGAGAAAGAAAGATTTA 3504
QY 3749 CTGCTTTTGAAGATGGAATCTAGTAAGGACACCTTACTGACTCAGCTGCTATAGG 3808
DB 3505 ACAGACAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3564
QY 3809 AATTTCCAGCTGAGAAATGGAATGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3868
DB 3565 GAACTGAATCTATTAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3624
QY 3869 AAAATATTTACTATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3928
DB 3625 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3684
QY 3929 CATGATGCTTAAACAAACAACTTACTGAAACATTTGACAGCTGATTAACACAGAT 3988
DB 3685 TTAATGAAAGATGCTGTTTAAATTAATGATGATGATGATGATGATGATGATGATGAT 3744
QY 3989 GAATTAAGGCGTCCAGAGAGAGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4048
DB 3745 GAGTTTAATGAAGTAAAGACAGATTTAATTAAGATGATGATGATGATGATGATGATGAT 3804
QY 4049 CAACAGCTTTAGATGTTCTGTAGAAATTCATCAACAAACATGATGATGATGATGATG 4108
DB 3805 AAGCATTAATCAAGATTTCTAAAGAAATTAATGATGATGATGATGATGATGATGATGAT 3857
QY 4109 CAGGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4168
DB 3858 -GAAAGAAAGTATGTAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3915
QY 4169 AGAGATGACCTTCAACATCTTGTAAAGCAATGATGATGATGATGATGATGATGATGATG 4228
DB 3916 TTAAGAGATGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3975
QY 4229 CATGTAATTTCTGTGAGGAGAAAGCTTGAATTCACAAAGAAAGAAAGAAAGAAAG 4288
DB 3976 GATATATTAAGAAAGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4035
QY 4289 AAGGATATTTGGGCAATCTGAAGAAAGCAAGTGTGAATGCTGTTGGAATCTA 4348
DB 4036 AAGGATTTA-----AAACATTTGAACAGATATTTTGAAGAGAAAGAAAGAAAGAAAG 4089
QY 4349 AAGGAAAGCAAAATTAAGCTCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4408
DB 4090 AAGGATCATTTTGAAGAAATTCAGAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4149
QY 4409 TTAAGTAAAGAGGTTCAGTGTTCACAGAAAGCTGTAAGATGATGATGATGATGATGAT 4468
DB 4150 AATTTAAAGAAAGTATCTTCAATTAAGAAAGTGAAGAAAGAAAGAAAGAAAGAAAGTAC 4209
QY 4469 AAGGACAGCTTAAGGCAAGCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4528
DB 4210 GAATTTAAAGAAAGGTAAGCAATTAATTAAGTGTGATGATGATGATGATGATGATGATG 4269
QY 4529 CTGTAACAAGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4588
DB 4270 GAAAGATGATTTGAAGAAATTAAGTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAG 4329
QY 4589 GACCATGAGAGAGACAGATTAAGTGTGCTGTAAGAAAGTGAAGTAAAGTCTTGA 4648
DB 4330 GATATGAAATTAAGGATTAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAG 4389
QY 4649 GAAAGAAAGTGAAGCAAGTATTAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAG 4708
DB 4390 GAAAGAAAGTGAAGTCTTAAAGATGTTTATGATGATGATGATGATGATGATGATGATG 4449
QY 4709 TCTAACAAGCTGAATCTCAGAAAGATGAGAGCTCCGTTCTATTAAGATGAC 4768
DB 4450 ATGAAAGAAAGAAAGAAAGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4509
QY 4769 CTGCAACAGAAAGCTGAAGAAAGCTGCTGATGATGATGATGATGATGATGATGATGATG 4828
DB 4510 GTTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4569


```
QY 4949 GAAACCTGTGAGATTGTGCTCCTTGAGAGAAAGATATTGTTAACTGAAGAA 5008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4931 GATATGATGAAATATAGGTGAGACAAAGATGAAGTTATAGTTAATAGTCAGAAA 4990
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5009 TTGCATCAAAAACTAATGACAGAAAAATTACTACATGAAAAAGATGAACCTGAACAA 5068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4991 GAGAAACGCATTGAAAAAGTTAAAGCGAAAAAGAAAAAATTAGAAAAAAGTTGAAGAA 5050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5069 GCTCAGGTGAGCTAAAGTGCAGGTGAGACATCTGATGAGAGTATGATCGAATCGAAG 5128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5051 GGTGTAGTGGTCTTAAAAAACACGTAGACGAGTAATGAAATATGTCAAAAAATTGAT 5110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5129 TCCTCACTGAGTCCCTTGACGATGAGAAACATGATGAAACACACTCTTGCTCTT 5188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5111 AAAGAAAGTTGATTAAGAAAGTATCTTAAGCTTTAGAAATCAAAAATGATGTTACTAATGTT 5170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5189 AAACAGCAGATGCAAGTACTCAAGAAAGAAAGAGCTGCAACAAACCATGAACAC 5248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5171 TTAACAAACAAATCAAGATTTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTAAAGTTA 5230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5249 TTAACAGC 5256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5231 TTTGCTGC 5238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: May 23, 2003, 07:54:33
Job time : 1348 secs

THIS PAGE BLANK (USPTO)